JUNE
1958
VOLUME 22
NUMBER 3

JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

JOURNAL OF CONSULTING PSYCHOLOGY

Edited by Lawrence F. Shaffer
Teachers College, Columbia University

Associate Editors: Donald W. Fiske, *University of Chicago.* Harrison G. Gough, *University of California.* A. Hunter [illegible], *University.* E. Lowell Kelly, *University of Michigan.* Morris Krugman, [illegible] Martha M. Luckey, *Board of Education, Cleveland, Ohio.* Fred [illegible], *New York, N. Y.* Carl R. Rogers, *University of Wisconsin.* [illegible] Edward Joseph Shoben, Jr., *Teachers College, Columbia.* [illegible] George J. [illegible], *University of Pittsburgh.* J. R. Wittenborn, *Western University.*

Arthur L. Hoffman, *Managing Editor*
Helen Orr, *Circulation Manager*
Barbara Cummings, *Editorial Assistant*

The *Journal of Consulting Psychology* is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy and counseling, personality, and the dynamics of behavior.

Editorial Office

Manuscripts submitted for *regular* publication should now be sent to the *Editor-elect*, Donald S. Zander, Department of Psychology, University of Michigan, Ann Arbor, Michigan.

Manuscripts submitted for *early* publication, Brief Reports, and tests for review should be sent, until August 1, 1956, to the Editor, Lawrence F. Shaffer, 525 West 120th St., New York 27, N. Y.

Style of Manuscripts. Manuscripts should be prepared in accordance with the *Publication Manual of the American Psychological Association, 1957 Revision*, obtainable for $1.50 from the Washington office of the Association.

Early Publication. Authors may request the early publication of an exceptional article in the first issue after it is passed by editorial readers by arranging for the author to pay the publication costs. Information about charges for early publication may be obtained from the Business Office.

American Documentation Institute. To save printing costs, materials supplementary to articles may be deposited with the ADI. To obtain copies, order by Document Number as cited in the reference footnote, from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. The indicated fee must be remitted in advance. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

Reprints. Authors receive fifty reprints gratis, except for early publication articles and Brief Reports. Additional quantities may be ordered when returning proofs.

Business Office

Subscriptions, advertising, subscriptions, change of address, claims for the nonreceipt of issues, and other business matters, should be sent to the American Psychological Association, Inc., [illegible] 16th St. N.W., Washington 6, D.C. Address changes must reach the Association [illegible] month of the month to take effect the following month. Undelivered copies of the Journal [illegible] changes will not be replaced; subscribers should notify the post office promptly and guarantee [illegible] second-class forwarding postage. Other claims for undelivered copies must be made within three months of publication.

The *Journal of Consulting Psychology* is published bimonthly in February, April, June, August, October, and December in one yearly volume comprising approximately 480 pages. The subscription per year [illegible] foreign postage, $6.50.

Published by the American Psychological Association, Inc.
at Lemon and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth St. N.W., Washington, D.C.

Entered as second-class matter [illegible] at the post office at Lancaster, Pa., under the act of March 3, 1879. Acceptance for mailing at special rate of postage provided for in Section 538, Act of February 28, 1925, authorized July 27, 1918. [illegible] the American Psychological Association, Inc.

JOURNAL OF CONSULTING PSYCHOLOGY

June, 1958 Vol. 22, No. 3

Contents

Intelligence in Childhood Schizophrenics, Other Emotionally Disturbed Children, and Their Mothers: *Anthony Davids* - 159
Some Personality Correlates of Subjects Who Fail to Appear for Experimental Appointments: *Allan H. Frey and Wesley C. Becker* - - - - - - - - - - - - - - - - 164
Normative Data and Factor Analysis on the Parental Attitude Research Instrument: *Marvin Zuckerman, Beatrice Barrett Ribback, Irwin Monashkin, and James A. Norton, Jr.* - 165
A Self-Scored Projective Measure of Achievement and Affiliation Motivation: *Richard deCharms* - 172
Asynchronism and Apparent Movement Thresholds in Brain-Injured Children: *Henry J. Mark and Benjamin Pasamanick* - 173
Opinions of Psychologists and Psychiatrists About Mental Health Problems: *Jum Nunnally* 178
Some Social-Psychological Characteristics of a Child Guidance Clinic Caseload: *Jack L. Roach, Orville Gurrslin, and Raymond G. Hunt* - - - - - - - - - - - - - - - 183
The Weighted *Q* Sort: A Procedure for Quantitatively Estimating Emotional Disturbance and Personality Change: *Marvin Reznikoff and Laura C. Toomey* - - - - - - - - 187
Adjustment and Prediction of Others' Behavior: *June Elizabeth Chance* - - - - - - - 191
Nonprojective Determinants of Perception on the TAT: *Bernard I. Murstein* - - - - 195
The Validity of the Objective Rorschach and the MMPI: *William D. Winter and Ramon A. Salcines* - 199
Age and Sex Influences on MMPI Profile Peak Distributions in an Abnormal Population: *Bernard S. Aaronson* - 203
Spearman Correlations Between MMPI Scores of College Students and Their Parents: *Andrew L. Sopchak* - 207
The MMPI in Cerebral Palsy: *Tom Linde and C. H. Patterson* - - - - - - - - - - 210
An Evaluation of the Comparability and Reliabilities of Two Behavior Rating Scales for Mental Patients: *Donald W. Stilson, Donald J. Mason, Malcolm D. Gynther, and Boris Gertz* - 213
The Relation of Difficulty in Drawing to Ratings of Adjustment Based on Human Figure Drawings: *Marvin J. Feldman and Raymond G. Hunt* - - - - - - - - - - - - - 217
On the Relationship Between Manifest Anxiety and Intelligence: *Charles D. Spielberger* - 220
Anxiety: Unitas or Multiplex?: *Douglas N. Jackson and Robert Bloomberg* - - - - - 225
Test Anxiety Questionnaires: *George Mandler and Judith E. Cowen* - - - - - - - - 228
Allergy and the Need for Recognition: *Norman S. Greenfield* - - - - - - - - - - - 230
Chance Frequency in Matching Problems When Success or Failure is Reported After Each Matching Operation: *Alan H. Roberts* - 233
An Unprofitable Application of the Semantic Differential: *Jack Block* - - - - - - - - 235
Varieties of Response Sets: *William E. Broen, Jr. and Robert D. Wirt* - - - - - - - - 237
Psychological Test Reviews - 241

Brief Reports

The *Journal of Consulting Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.

2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the American Documentation Institute to insure indefinite availability.

3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.

4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, including all matter except the title and the author's lines, must not exceed 75 lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style (1). Headings, tables, and references are avoided or, if essential, must be counted in the 75 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 75-line quota: [1]

[1] An extended report of this study may be obtained without charge from John Doe, 300 Market St., Prospect 6, Mass. (giving the author's full name and address), or for a fee from the American Documentation Institute. Order Document No. ——, remitting $—— for microfilm or $—— for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (*a*) The extended report should be typed with single spacing for economy in duplication. (*b*) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (*c*) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual* (1).

Reference

1. American Psychological Association. Council of Editors. *Publication manual of the American Psychological Association (1957 rev.).* Washington, D. C.: American Psychological Association, 1957.

Erratum

In the Brief Report by Walter F. McKeever and Alvin I. Gerstein, "Validity of the Hewson Ratios: Investigation of a Fundamental Methodological Consideration," this Journal, 1958, 22, 150, the second sentence in Footnote 1 should read, "Order Document No. 5467, remitting $1.75 for microfilm or $2.50 for photocopies."

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Intelligence in Childhood Schizophrenics, Other Emotionally Disturbed Children, and Their Mothers

Anthony Davids

Brown University and Emma Pendleton Bradley Hospital

Much of the research devoted to assessment of intellectual functioning in children and adults has been generated by the nature-nurture or heredity-environment controversy (1). In this regard, several previous studies have been concerned with intellectual similarities between (*a*) normal children and their parents, (*b*) normal siblings, (*c*) identical twins and fraternal twins, and (*d*) foster children and their foster parents. The effects on children's IQ of impoverished home environments and of institutionalization have been studied. And the effects of enriched environments and special schooling have also been related to intellectual functioning in children. Since excellent surveys of these numerous studies, with exhaustive bibliographies, are available elsewhere (1, 5, 11), they will not be given detailed consideration in this report.

In spite of the wealth of previous investigation in this topical area, there is still very little empirical evidence bearing on the problem of relations between intellectual functioning in severely disturbed children and in their mothers. In the attempt to gain further understanding in this regard, the present study seeks answers to the following four questions. Is there a significant correlation between the IQs obtained by children who are severely disturbed emotionally and the IQs obtained by their mothers? Is there a significant difference between the IQs of the mothers and the IQs of their disturbed children? Is there a significant difference between the IQs of mothers of schizophrenic children and the IQs of mothers of children suffering from other forms of emotional disturbance? Is there a significant difference between the IQs of schizophrenic children and the IQs of other types of emotionally disturbed children? The findings obtained in answer to these questions should be of interest both to investigators who are concerned with theoretical aspects of the problem of psychopathology and intellectual functioning, and to clinicians who are faced with the practical problem of building programs for assessing children and parents in the many new child treatment centers that are being established.

Method

The findings to be presented in this report are the result of a five-year program of psychological assessment of hospitalized emotionally disturbed children and their mothers.[1] The patients were 70 boys and girls, between the ages of 5 and 12, who were undergoing residential psychiatric treatment at the Emma Pendleton Bradley Home. Of this sample, 12 children were diagnosed as schizophrenic, and 58 were diagnosed as either severe behavior disorder or neurotic. Also serving as subjects in this study were the mothers of the children. These 70 women, between the ages of 26 and 46, were all living in the community. At the time of the child's admission to the institution, the mother was examined with a comprehensive battery of psychological tests, in-

[1] The assessment program was initiated by M. W. Laufer, E. Z. Rubin, and S. Koret. The author wishes to gratefully acknowledge the many hours of testing conducted by Rubin, M. Talmadge, Pearl Gore, and several psychology interns who have served at the Bradley Home. The conception of the present study, experimental design, statistical analyses of data, and this report are the contributions of the present author. The writing of this report was facilitated by Research Grants M-1722 and M-2087 from the National Institute of Mental Health, United States Public Health Service.

cluding the Wechsler Bellevue Intelligence Scale (short form), Rorschach test, TAT, and figure drawings. The children were also examined by experienced clinical psychologists who included the complete Wechsler Intelligence Scale for Children, along with projective tests and visual-motor tests, in their battery of techniques. The present report is concerned solely with the results of intelligence testing; the results of the mother-and-child personality assessment must await further work. Thus, the data for this study consist of the mothers' IQs based on their performance on the comprehension, vocabulary, picture completion, and block design subtests of the adult W-B, and the children's IQs based on their responses to the 10 subtests of the WISC.

Results and Discussion

The results show that there is a significant positive correlation ($r = .34$, $p = .01$) between the IQs of the 58 nonschizophrenic disturbed children and their mothers. In the schizophrenic group, however, there is no association ($r = .02$) between the indices of intelligence in the mothers and children. It should be pointed out that this correlation is based on the scores of only 8 schizophrenic children, as 4 of the group of 12 schizophrenics were untestable. They would neither attend to the test material nor attempt to cope with it, and in these cases only fragmentary, if any, test data were obtained, making any estimate of IQ practically impossible and even more unreliable than the estimates for the other schizophrenic children.

The coefficient obtained with the other types of disturbed children and their parents is lower than the coefficients in the vicinity of + .50 that have been reported (2, 3, 7, 8, 10) in studies of normal children and their parents, but even this degree of association is quite striking in view of the host of factors that would militate against finding a significant relation in our study. To mention only a few of these sources of error—the mothers' IQs were based on a shortened form of the intelligence test, the mothers were examined under conditions of stress (time of admission of child to hospital), and it is a well-known fact that emotionally disturbed children rarely are able to function at their optimal level on intelligence tests. Thus, in spite of these negative factors, a significant correlation was obtained, and one can only wonder what magnitude of relation might be found if these estimates of intelligence were not influenced by the intrusion of severe psychopathology.

Turning now to the differences between the performances of the mothers and children, presented in Table 1, a striking difference of 38 IQ points is found between the schizophrenics and their mothers, and a difference of 15 IQ points between the other disturbed children and their mothers. In both compari-

Table 1

Comparison of IQs of Mothers and of Their Emotionally Disturbed Children

Group	*N*	Mean	Variance	
Mothers of nonschizophrenics	58	108.1	175.8	
Mothers of schizophrenics	12	122.1	88.1	
Nonschizophrenic disturbed children	58	92.7	177.1	
Schizophrenic children	8	84.0	302.6	
Groups compared			*t*	*F*
Mothers of nonschizophrenics vs. mothers of schizophrenics			3.86**	2.00
Mothers of nonschizophrenics vs. their children			6.70**	1.01
Mothers of schizophrenics vs. their children			5.68**	3.44*
Nonschizophrenic disturbed children vs. schizophrenic children			1.65*	1.71

* Significant at .05 level.
** Significant at .001 level.

sons, these differences are significant well beyond the .001 level. Thus, it is evident that in the intellectual sphere the mothers fare much better than do their suffering offspring. It is clear that the burden of emotional conflict exacts a heavy intellectual toll in the young.

The two groups of mothers were compared. Even though the mean IQ of 108 obtained by the mothers of nonschizophrenic children is above the average for the general population, the mean IQ of 122 for the mothers of the schizophrenic children is significantly higher. This empirical finding is in keeping with clinical observations reported by several specialists in the field of childhood schizophrenia. Kanner, in describing the "indisputably high level of intelligence" of parents of cases of infantile autism, reports that "All but 5 of the mothers of the 55 children have attended college. All but one have been active vocationally before, and some also after, marriage as scientists, laboratory technicians, nurses, physicians, librarians, or artists. . . . My search for autistic children of unsophisticated parents has remained unsuccessful to date. . . . Most of the parents declare outright that they are not comfortable in the company of people; they prefer reading, writing, painting, making music, or just 'thinking' " (6, pp. 420–421). Rank, who also describes the use of intellectualization and obsessiveness in the parents of atypical children, reports that ". . . not too rarely they are highly intellectual, prominent people" (9, p. 131). Despert also emphasizes that mothers of these psychotic children seek gratification from intellectual sources rather than from contact with people. According to her, ". . . two significant characteristics of the mother and child which are frequently encountered in the group are, namely, the emotional detachment and the overintellectualization in the mother and the early absorption of the child in nonhuman content and nonfunctional activities, such as extensive acquisition of knowledge of musical records, license plate numbers, astronomical data, large and inappropriate vocabulary, word spelling, etc. On the one hand, the mother is remote and . . . on the other she puts excessive pressure on the growing baby to achieve feats of development which are considerably beyond his age level and interests" (4, p. 346).

The findings in the present research would, therefore, be predicted on the basis of these clinical reports. Discussion of similarities and differences between the present group of schizophrenics and those previously described as "autistic" or "atypical," as well as detailed description of the personalities of the mothers, are beyond the scope of the present report. It is only emphasized that for the kind of child diagnosed as schizophrenic at the Bradley Home, and for the kind of mothers of children who are admitted to this institution, the intelligence test findings fit well with previously reported observations.

Finally, the comparison between the IQs of the schizophrenic children and their nonschizophrenic emotionally disturbed peers who were also undergoing residential treatment is of interest. The schizophrenics, with an average IQ of only 84, score 8.7 IQ points below the other disturbed children. Since schizophrenic children are even more disturbed than the others, and since in most instances emotional disturbance detracts from intellectual functioning, this finding is in accord with the prediction of poorer performance by schizophrenics and, accordingly, is statistically significant.

In viewing these results of intelligence testing, it is noteworthy that the average IQ for this total sample is just about 92, which is considerably below the average for children in general. As mentioned above, some of the psychotic children are untestable, and as a group they perform very poorly. Even the ones who employ a mechanized, intellectualized approach to the world do not do very well on the standard tests because their specialized knowledge, obsessive concern for certain specific details, preoccupations with technical gadgets, astronomy, and so forth, does not earn them IQ points in the various areas of general knowledge that are tapped by conventional tests. Actually, the great store of anxiety with which they must cope and the kinds of defense mechanisms they employ make it impossible for them to benefit from contacts with people and events in their en-

vironment, and even the schizophrenic child who performs feats of memory in some particular area of concern or who seems to know a great deal about the physical universe or certain aspects of modern technology is unable to give an outstanding performance on intelligence tests. As for the over-all group of disturbed children, clinical psychologists who have tested them affirm that the majority are functioning considerably below their optimal level. Before being admitted to the Bradley Home, each child has to be evaluated as possessing the capacity for at least average intellectual functioning. That is, the psychologist must judge that even though the child tests subnormal (sometimes falling even into the "mental defective" category), basically, he is of normal or higher intelligence. The fact that as a group their intellectual functioning is greatly impaired is further demonstrated by the finding, mentioned above, that they perform so much worse than do their mothers. It seems unlikely that such discrepancies would be found between groups of normal children and their parents. In fact it seems highly plausible that, in many situations, groups of normal children would obtain higher IQs than their parents.

It is hoped that future studies of the personalities of these children and their mothers will contribute to further understanding of the dynamics of the psychopathology and nature of the intellectual deficit. It is evident from the present findings, however, that severe emotional disturbance has a pronounced detrimental effect on intellectual functioning in children. Undoubtedly, this state of affairs must set up a vicious circle, with disturbance leading to increasing intellectual failure and frustration, and the inability to achieve intellectual mastery leading to greater emotional maladjustment. In a residential treatment center, the therapeutic approach is from both points; with child care workers and recreational personnel helping the child to acquire increasing skills and abilities, special school teachers providing individualized and understanding teaching, and psychotherapists working with the child's emotional conflicts. The goal, of course, is to acquire emotional and intellectual light where now there is darkness. Objective knowledge of the intellectual and emotional fruits to be derived from these therapeutic efforts awaits much further investigation, but there is no doubt that these children are deserving of the scientific and clinical investment.

Summary

The purpose of this investigation was to study the intellectual functioning of 70 hospitalized emotionally disturbed children, of whom 12 were schizophrenic and 58 were suffering from other types of severe disturbance, and to compare the children's intelligence with a measure of their mothers' intelligence based on performance on the Wechsler-Bellevue scales. A significant positive correlation was found between the IQs of nonschizophrenic children and the IQs of their mothers, but no association was found between the IQs of the schizophrenic children and their mothers. Both groups of children received IQs significantly lower than did the groups of mothers. Although the mothers of the nonschizophrenic children received a mean IQ that was higher than the mean for the general population, the mean IQ of 122 for mothers of schizophrenic children was significantly higher. In comparing the two groups of children, it was found that the schizophrenics received significantly lower IQs, and that the mean IQ for this total sample of disturbed children fell at the bottom of the normal range of intelligence. Some theoretical, therapeutic, and practical implications of these findings were discussed.

Received June 27, 1957.

References

1. Anastasi, Anne, & Foley, J. P. *Differential psychology*. New York: Macmillan, 1949.
2. Burks, Barbara S. The relative influence of nature and nurture upon mental development: A comparative study of foster parent-foster child resemblance and true parent-true child resemblance. *Yearb. nat. Soc. Stud. Educ.*, 1928, **27**, Part I. Pp. 219–236.
3. Conrad, H. S., & Jones, H. E. A second study of familial resemblance in intelligence: environmental and genetic implications of parent-child and sibling correlations in the total sample. *Yearb. nat. Soc. Stud. Educ.*, 1940, **39**, Part II. Pp. 97–141.

4. Despert, J. Louise. Some considerations relating to the genesis of autistic behavior in children. *Amer. J. Orthopsychiat.*, 1951, **21**, 335–347.
5. Jones, H. E. The environment and mental development. In L. Carmichael (Ed.), *Manual of childd psychology*. New York: Wiley, 1952. Pp. 631–696.
6. Kanner, L. Problems of nosology and psychodynamics of early infantile autism. *Amer. J. Orthopsychiat.*, 1949, **19**, 416–426.
7. Leahy, A. M. Nature-nurture and intelligence. *Genet. psychol. Monogr.*, 1935, **17**, 236–308.
8. Pearson, K. *Nature and nurture*. London: Dulau, 1910.
9. Rank, Beata. Adaptation of psychoanalytic technique for the treatment of young children with atypical development. *Amer. J. Orthopsychiat.*, 1949, **19**, 130–139.
10. Roff, M. A statistical study of the development of intelligence test performance. *J. Psychol.*, 1941, **11**, 371–386.
11. Thompson, G. G. *Child psychology*. Boston: Houghton Mifflin, 1952.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Some Personality Correlates of Subjects Who Fail to Appear for Experimental Appointments[1]

Allan H. Frey and Wesley C. Becker
University of Illinois

In the course of an experiment using two extreme groups, introverts and extraverts, the characteristics of *Ss* who failed to appear for an appointment with or without notifying the experimenter were investigated.

Five of the Guilford personality scales, Sociability, Rathymia, General activity, Social ascendance, and Emotional stability, were administered to a large section (250) of introductory psychology students. These tests were scored, and a group of introverts and a group of extraverts were selected by the experimenter on the basis of extreme Rathymia scores. Appointments were then made for each member of these groups by the use of postal cards.

Twelve *Ss* notified the experimenter that they would be unable to keep the appointment (Group *N*); fourteen *Ss* failed to appear without notifying the experimenter (Group *F*). The *N* group was characterized by low scores on scales *S*, *R*, *G*, and *A*; the *F* group by high scores. The chi-square test indicated that these differences are significant at the .01 level, except *R*, which is significant at the .001 level. There was no significant difference between groups on the *C* scale, nor was there a significant difference within or between groups with regard to sex. According to Eysenck, the *S*, *R*, *G*, and *A* scales are all highly loaded on a second-order factor which has been called extraversion–introversion.

The results of this study indicate that a sampling bias occurs when *Ss* are obtained by the method noted above. This bias could have undesirable effects in many kinds of experiments. For example, in several recent studies, Eysenck has been attempting to relate extraversion–introversion to such variables as degree of satiation, conditionability, and reminiscence in pursuit rotor learning. Sampling bias of the kind indicated above would tend to attenuate any true relationships between these variables and extraversion–introversion. Of course, the generalizability of any findings is also affected. It is suspected that such bias would also enter most studies of conformity behavior, group dynamics, and anxiety. Anxiety studies enter the picture because of their correlation with extraversion–introversion. Franks reported an r of $-.48$ between Rathymia and the Taylor Anxiety scale. Farber *et al.* reported an r of $-.36$ between the same two tests. The present writers found a correlation of $-.32$ between Cattell's Anxiety scale and Eysenck's Extraversion scale. It therefore seems likely that selective sampling on extraversion would encompass to a lesser degree selective sampling on anxiety.

In practice, we have found it possible to reschedule the *N* subjects, but unless considerable pressure is exerted on the *F* subjects, they rarely appear when rescheduled. In experiments where extraversion–introversion is suspected as an important factor, it is suggested that pretesting on this variable be carried out while *Ss* are in a "captive" state, e.g., in the classroom, and that special pressure be exerted to obtain cooperation from the extreme group members.

Brief Report.
Received January 2, 1958.

[1] An extended report of this study may be obtained without charge from Allan H. Frey, Department of Psychology, University of Illinois, Urbana, Ill., or for a fee from the American Documentation Institute. Order Document No. 5537, remitting $1.25 for microfilm or $1.25 for photocopies.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Normative Data and Factor Analysis on the Parental Attitude Research Instrument[1]

Marvin Zuckerman and Beatrice Barrett Ribback
Indiana University Medical Center

Irwin Monashkin and James A. Norton, Jr.
Purdue University

The Parental Attitude Research Instrument (PARI), developed by Schaefer and Bell (3), consists of 23 five-item scales. Each scale measures an attitude which is theoretically relevant to personality development in children. The test items are modeled after the type of items developed by Mark (2) and Shoben (4) in their studies of parental attitudes. The items are general opinions with which the subject is asked to strongly or mildly agree or disagree on a four-point scale. An attempt was made to increase the social acceptability of items by putting them in the form of rationalizations, truisms, and cliches. The names of the 23 scales can be found in Table 4.

Schaefer and Bell's normative multiparae sample was 100 wives of military personnel who were given the test one to four days following delivery at a military hospital. There are two limitations to the representativeness of this sample: (*a*) the selectness of the occupational group and (*b*) the possibility that the expressed parental attitudes may have been influenced by the recency of parturition.

The senior author is engaged in a series of studies on parental attitudes in various groups of women including mothers of disturbed children. In collecting a normative sample we had four purposes: (*a*) to provide stanine scores for future work, (*b*) to investigate four possible influences on parental attitudes: age, education, number of children, and religion, (*c*) to obtain data for a factor analysis of the 23 scales, and (*d*) to provide a pool of control subjects for comparisons with other groups of mothers.

In order to facilitate interpretation in comparisons between different groups of mothers, it was deemed desirable to determine the factorial structure of the test. Schaefer and Bell analyzed the factorial structure of the test on their normative group of 100 mothers using the grouping method of factoring without rotation of axes. Their results revealed three factors in the test: Factor I included 16 of the 23 scales and seemed to be a controlling-authoritarian factor; Factor II included four scales and seemed to be a hostility-rejection factor; Factor III included three scales and seemed to be a democratic-equalitarian factor. The purpose of our factor analysis was to derive a single factorial structure from heterogeneous groups which would be useful in comparing factor scores across such groups; and to see if the factorial structure derived from such a heterogeneous sample would be similar to that obtained by Schaefer and Bell from a simple normal sample.

Procedure

Normative Data

The normative sample was composed of 222 mothers, including mothers of nursery school children, mothers of young adults attending a night extension college, and mothers

[1] The authors would like to thank Donald F. Moore, Superintendent of Carter Memorial Hospital, for his kind permission to use the facilities of Carter Hospital to collect data from patients and their mothers. A note of thanks is extended to James Besheers for his assistance in the mechanics of the factor analysis.

contacted at meetings of church social groups. Relevant data on this sample appears in Tables 1–3. Using church social groups gave us the advantage of being able to secure "captive" groups, thus avoiding the selective effects entailed in using volunteers. However, use of these groups may introduce another selective factor since they comprise slightly over one-half of the normative sample. In addition, it should be noted that the sample is drawn exclusively from an urban, midwestern population. It contains a greater representation of women with some college education and a greater proportion of older mothers than is found in the general population. The religious and racial proportions of the sample are not too deviant from those of the general population.

Factor Analysis

The following subjects were added to the normal group of 222 to comprise the sample for the factor analysis: (*a*) 60 mothers who were psychiatric patients in an acute treatment hospital, (*b*) 69 mothers of young adult psychiatric inpatients, and (*c*) 62 mothers of children referred for examination to two child guidance clinics and a children's inpatient psychiatric service. The total *N* for the factor analysis was therefore 413. The mean education of this sample was 11.9 years, *SD* 2.9. The mean age was 38.6, *SD* 10.1.

Scores of all 413 subjects were punched on cards and intercorrelated on the 650 IBM machine at Indiana University. The largest absolute value of the correlations in each column of the matrix was used as the communality estimate for the scale. The matrix was factored, using the principal axes solution, on Purdue University's Datatron computer. The first three principal components accounted for a total of 11.96 out of an estimated communality variance of 13.57. It was decided, tentatively, to stop with three factors, since the principal components after the third (which were also computed) added only

Table 1

Education Distribution of Normative Sample
(M = 12.73 SD = 2.79)

Grades	*N*	%
0–6	4	2
7	5	2
8	13	6
9	6	2
10	6	2
11	9	4
12	79	36
13	17	8
14	22	10
15	11	5
16	44	20
17	2	1
18	2	1
19	1	1
20	1	1
	222	99

Table 2

Age Distribution of Normative Sample
(M = 37.94 SD = 8.71)

Age	*N*	%
0–19	1	0
20–24	5	2
25–29	25	11
30–34	59	27
35–39	47	21
40–44	40	18
45–49	24	11
50–54	9	4
55–59	7	3
60–64	3	1
65–70	2	1
	222	98

Table 3

Number of Children, Distribution in Normative Sample
(M = 2.53 SD = 1.46)

No.	*N*	%
0	6	3
1	46	21
2	76	34
3	48	22
4	28	12
5	5	2
6	11	5
8	1	0
9	1	0
	222	99

Table 4

Means, *SD*s, and Stanine Scores for the PARI
($N = 222$)

Scale		9	8	7	6	5	4	3	2	1	Mean	*SD*
1	Encouraging Verbalization	5-14	15	16	17	18	19	20	—	—	17.93	1.95
2	Fostering Dependency	20-17	16-15	14	13-12	11	10-9	8-7	6	5	10.82	3.08
3	Seclusion of the Mother	20-18	17-16	15-14	13	12-11	10-9	8	7-6	5	11.39	3.38
4	Breaking the Will	20-17	16-15	14	13-12	11-10	9	8-7	6-5	—	10.54	3.29
5	Martyrdom	20-16	15-14	13	12-11	10-9	8-7	6	5	—	9.44	3.54
6	Fear of Harming the Baby	20-19	18	17-16	15-14	13-12	11	10-9	8-7	6-5	12.81	3.40
7	Marital Conflict	—	20	19-18	17-16	15	14-13	12-11	10	9-5	15.04	3.33
8	Strictness	20	19-18	17	16-15	14-13	12	11-10	9	8-5	13.71	3.22
9	Irritability	20	19-18	17	16-15	14	13	12-11	10	9-5	14.22	2.81
10	Excluding Outside Influences	20-18	17-16	15-14	13-12	11	10-9	8-7	6-5	—	10.90	3.54
11	Deification	20	19-18	17-16	15-14	13-12	11	10-9	8-7	6-5	12.80	3.69
12	Suppression of Aggression	20-15	14	13	12-11	10	9-8	7	6	5	10.03	2.75
13	Rejection of Homemaking Role	20-18	17-16	15	14-13	12-11	10	9-8	7-6	5	11.62	3.23
14	Equalitarianism	5-13	14	15	16	17	18	19	20	—	17.40	2.03
15	Approval of Activity	20-19	18-17	16	15-14	13	12-11	10	9-8	7-5	12.84	3.02
16	Avoidance of Communication	20-17	16-15	14	13-12	11	10-9	8-7	6	5	10.81	3.20
17	Inconsiderateness of Husband	20-19	18-17	16-15	14-13	12	11-10	9-8	7-6	5	11.96	3.61
18	Suppression of Sex	20-14	13	12-11	10	9-8	7	6-5	—	—	8.71	2.97
19	Ascendance of the Mother	20-19	18-17	16-15	14-13	12-11	10	9-8	7-6	5	11.83	3.55
20	Intrusiveness	20-17	16-15	14	13-12	11-10	9-8	7-6	5	—	10.39	3.51
21	Comradeship and Sharing	5-15	16	17	18	19	—	20	—	—	18.67	1.70
22	Acceleration of Development	20-18	17	16-15	14-13	12-11	10-9	8	7-6	5	11.67	3.56
23	Dependency of the Mother	20-18	17	16-15	14-13	12	11-10	9-8	7	6-5	12.00	3.24

small amounts to the total of explained variance. This decision was confirmed by examination of the residuals after the three factors: of 276 residuals, only seven were over .10, the largest being .16. It was felt that this was well within the sampling variation of the original correlations (for 400 *df*, the 5% level for significance of difference is .098). Orthogonal rotations were carried out on the factor matrix using the single plane method (actually only one rotation in the plane of the second and third components seemed indicated).

Results and Discussion

Normative Data

The scales and their means, standard deviations and stanine scores derived from the normal group, are given in Table 4. The correlations of the scales with age, education, and number of children are given in Table 5. The education of the mother seems to be the most significant variable of the three, relating to parental attitudes. Using the one per cent level for significance, education correlates significantly with 18 of the 23 scales, age with six of the scales, and number of children with none of the scales. The direction of the correlations indicate that the less educated and older mothers tend to have more severe attitudes (e.g., more authoritarian, suppressive, hostile) on the scales which are related.

Samples of 38, equivalent in age and education, were taken from each of two religious groups, Catholic and Protestant. Comparisons of the means on the individual scales were made using simple analysis of variance. A significant difference was found only on Scale 7, Marital Conflict, on which the Catholics scored significantly higher. However, one difference out of 20 tests might be expected by chance; therefore, we must not conclude that religion (Catholic and Protestant) is a significant influence on parental attitudes as measured by the PARI. It might be noted, however, that there exists a rather homogeneous culture in this area. Differences between religious groups may be greater in sections of the country where there is less homogeneity of culture. The size of the samples

also limits any broad conclusions drawn from these data.

Factor Analysis

The correlation matrix of the 23 PARI scales, education and age, based on the total *N* of 413, can be seen in Table 6. The figures in the diagonals are the communality estimates. The unrotated factor matrix can be seen in Table 7 and the rotated factor matrix in Table 8.

The three factors extracted in this study are very similar to those found by Schaefer and Bell, with some minor exceptions which will be noted in the discussion of Factor B. Thus, the factor structure over a clinically heterogeneous group of mothers appears to be substantially the same as that obtained on normal mothers only.

Factor A has loadings of .60 and above on 16 of the 23 scales. The significant loadings are on the same scales as found in Schaefer and Bell's Factor I. The loadings are generally higher in this study, probably reflecting the greater heterogeneity of our sample. The scales which contribute most heavily to this factor are those which measure authoritarian, suppressive, punitive, and restricting types of attitudes. However, this factor is so broad that it may be measuring a general response set to the test as well as specific response to the content of the scales. Education has a negative loading on this factor indicating that the more highly educated mothers are more likely to score low on this factor.

Factor B is similar to Factor II in Schaefer and Bell's analysis except that Scale 8, Strictness, no longer is part of the factor. Factor B is defined mainly by the two scales, 9 and 13, Irritability and Rejection of the Homemaking Role. Scale 7, Marital Conflict, shows a smaller loading on this factor. The character of this factor can be described as hostility toward children and husband and rejection of the maternal role. Unlike Factor A, this fac-

Table 5

Correlations Between PARI Scales and Age, Education, and Number of Children*

Scale		Education	Age	No. of children
1	Encouraging Verbalization	.15	.05	−.20
2	Fostering Dependency	−.34*	.11	.07
3	Seclusion of the Mother	−.47*	.14	.19
4	Breaking the Will	−.41*	.13	.14
5	Martyrdom	−.41*	.09	.04
6	Fear of Harming the Baby	−.31*	.08	−.03
7	Marital Conflict	−.21*	.04	.18
8	Strictness	−.25*	.17*	.11
9	Irritability	−.08	−.14	.23
10	Excluding Outside Influences	−.46*	.16	.19
11	Deification	−.37*	.01	.18
12	Suppression of Aggression	−.27*	.02	.07
13	Rejection of Homemaking Role	−.06	−.04	.03
14	Equalitarianism	.12	.11	−.05
15	Approval of Activity	−.26*	.17*	.10
16	Avoidance of Communication	−.39*	.29*	.02
17	Inconsiderateness of Husband	−.29*	.18*	.01
18	Suppression of Sex	−.43*	.22*	.00
19	Ascendance of the Mother	−.32*	.12	.08
20	Intrusiveness	−.31*	.00	.00
21	Comradeship and Sharing	−.01	.02	.10
22	Acceleration of Development	−.42*	.23*	.10
23	Dependency of the Mother	−.29*	.01	.00

* Significant at or below the .01 level.
* Number of children was not available for part of the normative sample at the time the correlations were run. The correlations given for age and education are based on an *N* of 222. Correlations for No. of Children are based on an *N* of 125.

Table 6

Correlation Matrix; 23 PARI Scales, Education, and Age

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Ed.	Age
1	(36)	−07	−13	−13	−19	−05	04	−10	06	−21	−18	−11	00	36	−07	−15	−10	−23	−12	−20	28	−13	−08	13	01
2		(59)	55	53	57	46	25	33	05	59	53	52	15	02	44	43	40	58	49	55	13	55	50	−45	25
3			(62)	54	59	50	31	39	06	62	50	48	12	11	52	47	51	62	56	55	08	60	47	−53	20
4				(65)	65	45	36	48	26	62	60	48	27	01	49	54	46	59	54	55	10	56	48	−47	18
5					(65)	50	32	40	19	60	55	53	27	−03	48	49	50	64	53	58	00	53	51	−44	13
6						(53)	31	26	32	50	42	38	35	16	45	46	43	53	51	46	10	46	51	−35	10
7							(38)	23	33	24	23	10	26	16	30	29	38	20	31	17	12	28	26	−22	01
8								(50)	15	37	48	29	16	08	50	38	27	33	43	40	19	38	34	−21	20
9									(51)	09	12	03	51	15	12	24	21	−01	18	01	06	05	23	02	−08
10										(70)	70	57	16	02	50	58	48	65	54	62	10	56	53	−51	20
11											(70)	48	09	−03	48	49	35	58	53	59	18	52	48	−41	14
12												(57)	22	−01	43	44	37	57	40	53	04	47	44	−39	19
13													(51)	17	20	27	34	12	26	12	05	14	31	−03	02
14														(37)	15	05	11	−06	09	−04	37	09	08	06	02
15															(59)	56	46	47	59	55	14	54	44	−31	22
16																(58)	43	54	50	48	05	49	40	−37	34
17																	(51)	44	49	38	04	49	47	−32	21
18																		(65)	55	60	−01	61	49	−50	29
19																			(59)	57	14	46	42	−39	21
20																				(62)	08	55	46	−37	12
21																					(37)	13	11	−05	06
22																						(61)	49	−48	32
23																							(53)	−33	10
Ed.																								(53)	−23
Age																									(34)

Table 7

Unrotated Factor Matrix

Scale No.	Name	Factor I	Factor II	Factor III	h^2
1	Encouraging Verbalization	−.18	.37	−.40	.33
2	Fostering Dependency	.71	−.10	−.11	.53
3	Seclusion of the Mother	.76	−.08	−.10	.59
4	Breaking the Will	.77	.04	.08	.61
5	Martyrdom	.77	−.04	.18	.63
6	Fear of Harming the Baby	.66	.22	.09	.48
7	Marital Conflict	.39	.36	.07	.29
8	Strictness	.54	.09	−.13	.31
9	Irritability	.21	.60	.34	.52
10	Excluding Outside Influences	.80	−.15	−.01	.67
11	Deification	.74	−.12	−.06	.56
12	Suppression of Aggression	.65	−.15	.00	.45
13	Rejection of the Homemaking Role	.31	.53	.33	.48
14	Equalitarianism	.07	.49	−.40	.40
15	Approval of Activity	.70	.10	−.13	.51
16	Avoidance of Communication	.70	.06	.05	.49
17	Inconsiderateness of the Husband	.63	.17	.10	.43
18	Suppression of Sex	.78	−.26	.03	.67
19	Ascendance of the Mother	.72	.08	−.02	.53
20	Intrusiveness	.73	−.19	−.02	.58
21	Comradeship and Sharing	.14	.31	−.49	.35
22	Acceleration of Development	.74	−.08	−.14	.57
23	Dependency of the Mother	.66	.13	.07	.45
24	Education	−.58	.22	.07	.38
25	Age	.28	−.11	−.22	.14
					11.97
Sum Aij^2		9.37	1.61	0.98	11.96

tor seems to be unrelated to the education of the mother.

Factor C is identical to Factor III in the Schaefer and Bell analysis and consists of three scales: 1, Encouraging Verbalization; 14, Equalitarianism; and 21, Comradeship and Sharing. Factor C can be described as measuring democratic attitudes toward child rearing. Theoretically, one would not expect to find this as an independent factor since these scales should load on the negative side of Factor A, Authoritarianism. Apparently, some mothers score high on both of these factors which are somewhat contradictory in content. Perhaps the acquiescent response set, discovered in the F scale by Bass (1), may be responsible for this result. Many mothers may agree with contradictory statements because of the tendency to agree with any sort of positive sounding generality. Acquiescent response set on the PARI is currently being investigated by the senior author.[2]

Summary

Normative data on the Parental Attitude Research Instrument were gathered from 222 mothers. Stanine scores for the 23 scales were derived from the data. Correlations were done between the scales and the education, age, and

[2] Subsequent to this factor analysis of the total heterogeneous sample, separate factor analyses were carried out on the group of 131 mothers of disturbed children and adults and on the group of 222 mothers of normals. Although the correlations between scales were higher in the group of mothers of disturbed children and adults than in the mothers of normals, Factors A, B, and C emerged from each analysis and were defined by essentially the same scales as reported in the analysis of the total sample. These results indicate that the obtained factors are stable over two different samples and may be used for comparisons between similar criterion groups.

Table 8

Rotated Factor Matrix

Scale No.	Name	A	B	C	h^2
1	Encouraging Verbalization	−.18	.07	.55	.34
2	Fostering Dependency	.71	−.15	.03	.53
3	Seclusion of the Mother	.76	−.12	.03	.59
4	Breaking the Will	.77	.08	−.04	.61
5	Martyrdom	.77	.07	−.17	.63
6	Fear of Harming the Baby	.66	.23	.05	.48
7	Marital Conflict	.39	.34	.15	.29
8	Strictness	.54	.00	.15	.31
9	Irritability	.21	.69	.07	.52
10	Excluding Outside Influences	.80	−.12	−.08	.67
11	Deification	.74	−.14	−.02	.56
12	Suppression of Aggression	.65	−.12	−.09	.45
13	Rejection of the Homemaking Role	.31	.62	.03	.48
14	Equalitarianism	.07	.17	.60	.40
15	Approval of Activity	.70	.00	.16	.51
16	Avoidance of Communication	.70	.08	−.01	.49
17	Inconsiderateness of the Husband	.63	.20	.02	.43
18	Suppression of Sex	.78	−.20	−.17	.67
19	Ascendance of the Mother	.72	.06	.07	.53
20	Intrusiveness	.73	−.17	−.09	.58
21	Comradeship and Sharing	.14	−.03	.58	.35
22	Acceleration of Development	.74	−.15	.07	.57
23	Dependency of the Mother	.66	.14	.01	.45
24	Education	−.58	.21	.07	.38
25	Age	.28	−.21	.11	.14
					11.97
Sum Aij^2		9.37	1.40	1.20	11.97

number of children of the mother. The education of the mother was most significantly related to scores on the scales. Age showed a smaller number of relationships, and number of children no significant relationships with the scales.

The scales were factor analyzed using a heterogeneous sample composed of the 222 mothers in the normative sample, plus 131 mothers of psychiatric patients and disturbed children and 60 mothers who were psychiatric patients. Three factors were extracted from the test:

1. Factor A, Authoritarian-Control
2. Factor B, Hostility-Rejection
3. Factor C, Democratic Attitudes

The influence of response set on the test was considered.

Received July 1, 1957.

References

1. Bass, B. M. Authoritarianism or acquiescence? *J. abnorm. soc. Psychol.*, 1955, **51**, 616–623.
2. Mark, J. C. The attitudes of the mothers of male schizophrenics toward child behavior. *J. abnorm. soc. Psychol.*, 1953, **48**, 185–189.
3. Schaefer, E. S., & Bell, R. Q. Development of the Parental Attitude Research Instrument. *Child Developm.*, in press.
4. Shoben, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, **39**, 101–148.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

A Self-Scored Projective Measure of Achievement and Affiliation Motivation[1]

Richard deCharms
Washington University

The measure of motivation recently developed by McClelland *et al.* (1) for experimental purposes has the disadvantage of requiring a careful content analysis of projective materials. The procedure to be described here was developed with the aim of obtaining quickly a motive score that is as close to the score derived from content analysis as possible.

Each *S* wrote six stories in response to six picture stimuli under the standardized conditions described by McClelland *et al.* (1). *S* was then given a form containing several sample story plots for each stimulus picture and asked to rank order their similarity to his own story. A score was derived from this form by totalling the rankings of motive relevant plots on the latter form. These "self-scored" motive scores were compared with the scores derived from content analysis of the original stories written by *S*.

Achievement motivation. Scores for three groups of *S*s (total $N = 183$), were obtained from the self-scored measure by totalling the rankings of achievement related plots. Analysis of variance of the means of the content analysis measure classified as high, medium, and low by the self-scored measure demonstrated that the means differed significantly in all three groups ($p < .001$). For a fourth group ($N = 47$), the self-scored measure of achievement motivation discriminated significantly between volunteers and nonvolunteers for a psychological experiment. The content analysis score has previously been shown to be related to volunteering.

Affiliation motivation. Scores for three groups (total $N = 128$) were obtained from the self-scored measure by totalling the rankings of the affiliation related plots. Analysis of variance of the means of the content analysis measure classified as high, medium, and low by the self-scored affiliation measure demonstrated that the means differed significantly in all three groups ($p < .025$).

The self-scoring procedure should not be as sensitive and reliable as actual content analysis. It will, however, yield an immediate score which could be used to identify *S*s with high or low motive scores, and will be especially valuable where the test could be administered to a large number of *S*s from which extremes could be drawn. Since much of the work of scoring has been eliminated, the new measure will allow the researcher to use more *S*s or to replicate results.

Brief Report.
Received December 14, 1957.

[1] An extended report of this study may be obtained without charge from Richard deCharms, Graduate Institute of Education, Washington University, St. Louis 5, Missouri, or for a fee from the American Documentation Institute. Order Document No. 5536, remitting $1.75 for microfilm or $2.50 for photocopies.

Reference

1. McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. *The achievement motive.* New York: Appleton-Century-Crofts, 1953.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Asynchronism and Apparent Movement Thresholds in Brain-Injured Children[1]

Henry J. Mark
Department of Otolaryngology, Johns Hopkins University

and Benjamin Pasamanick
Department of Psychiatry, Ohio State University

This paper reports some differences in foveal asynchronism and apparent movement thresholds, and peripheral two-point and apparent movement thresholds, found between brain-injured (pyramidal tract damaged) and control children. The results were obtained in the course of a larger study aimed at discovering neuropsychological measurements which might be used to detect otherwise unnoticed alterations in cerebral functioning. Previous studies by Klein and Krech (4) and Eysenck (3) demonstrated altered figural after-effects in brain injury, and studies with the same subjects (*Ss*) as those reported here with critical flicker fusion (CFF) (8) and light perception (7) threshold measurements suggest that visual threshold measurements may be sensitive to the effect of nonoccipital lobe lesions.

Asynchronism threshold measurements were selected for study here to determine whether lowered visual efficiency and increased intra-individual variability in the brain-injured could again be demonstrated as in the previous studies with these groups (7, 8). Apparent movement threshold measurements were selected for study because an explanation of the phenomenon of apparent movement usually invokes one form or another of a stimulus after-effect or stimulus surround-effect theory. By presenting binocular stimuli in homonymous half fields, Birch and Ashdourian (2) demonstrated that central mechanisms by themselves are capable of producing an apparent movement (phi) phenomenon. Since central mechanisms are involved, such mechanisms may then give rise to differences in apparent movement thresholds between brain-injured and control groups. In addition, peripheral mechanisms such as retinal facilitation (9, p. 140), which are known to contribute to the perception of the apparent movement phenomenon, may well be altered in children whose brain injury has been present since birth. This may be hypothesized on the basis of epidemiologic studies (5, 6, 10, 11, 12) which point to anoxia (rather than traumatic injuries) as the major cause of brain damage. Since the neurons of the retina are also susceptible to damage from anoxia, altered retinal stimulus after-effects or surround-effects may combine with central mechanisms to produce apparent movement thresholds which distinguish brain-injured from control *Ss*.

Hypotheses

Higher asynchronism thresholds, corresponding to loss of visual efficiency in the detection of a pause (p) between the onset time of two lights, were predicted in the brain-injured. With p as the variable, apparent movement thresholds and peripheral two-point thresholds cannot be readily interpreted in terms of

[1] This research was supported by the Alfred P. Sloan Foundation, Inc. grant to the Division of Laryngology and Otology, Department of Surgery, Johns Hopkins University, as well as by a grant from the Foundation for Mentally Retarded and Handicapped Children of Baltimore. The authors wish to thank Paul Meier for invaluable assistance and advice, and both Bernard Weiss and Victor Laties for their critical reading of the manuscript.

visual efficiency, and the direction of the difference between brain-injured and control *Ss* in these thresholds was not specified in our hypothesis. Greater intra-individual variability, reported as the *V* score, was predicted in the brain-injured. It was also hoped that differential reactions to peripheral as compared to foveal measurements would again provide good discriminators between brain-injured and control *Ss*, as it did in two CFF experiments (7, 8).

The Experiment

Subjects

The brain-injured group was composed of ten children (six males, four females) with diagnosed pyramidal tract involvement and with no primary diagnosed opthalmologic pathology. The control group was composed of ten children (five males, five females) with conditions which did not justify an assumption of brain damage or injury such as spina bifida, hemophilia, or congenital amputation. All *Ss* attended a Baltimore school for handicapped children. The groups were roughly comparable in terms of age, IQ, and sex, and the effects of these covariates on all significant discriminators were examined by means of regression diagrams. None of these was significantly related to performance. For example, there was no apparent regression of thresholds or *V* scores on IQ (which ranged between 64 and 121), and it was therefore assumed that IQ did not enter as an important experimental factor in the difference between brain-injured and control groups. The mean IQ of the brain-injured group was 89, with an *SD* of 16; the mean for the control group was 96, with an *SD* of 16. The age range was between 10 and 15 years, with a median of 12 for both groups.

Apparatus

Two Sylvania R1131 C glow-modulator tubes were mounted centrally on a perimeter so that two rectangular focused lines of light, each $\frac{3}{4}$-mm wide and 6-mm long, and separated by a distance of 4 mm, could be projected by means of a simple achromatic lens onto a piece of white plastic mounted on black velvet cloth. The glow modulators were activated by an electronic pulse-type stimulator and timer (13) which permitted continuous variation from 4 ms to 260 ms, of the pause (p) between the onset of the two light stimuli. The lights were viewed by the *Ss* from a distance of 33 cm. The two lights were matched for wave length and had a brightness of the order of one-tenth millilambert (0.1 ml) as determined with a MacBeth Illuminometer. The light stimuli were presented once per second with the duration of both stimulus lights held constant at 135 ms. For peripheral threshold determinations, a 5-mm diameter circle of luminous paint served as a fixation point. The experiments were carried out in total darkness.

Procedure

The first binocular foveal measurement was always an ascending asynchronism threshold (A_{Asc}) measurement, where the pause (p) between the onset of the two light stimuli was continuously increased from 4 ms until *S* reported that the two lights were no longer presented simultaneously. As p was further increased, all *Ss* noted movement effects, and the ascending apparent movement threshold (M_{Asc}) was recorded when *S* reported that the observed phenomenon appeared equivalent to the real movement of one line moving from left to right. Similarly, descending apparent movement (M_{Desc}) and asynchronism threshold (A_{Desc}) measurements, which may be considered simultaneity thresholds, were obtained. In this way, six ascending and six descending threshold measurements were obtained foveally. These were followed by three left-eye peripheral threshold measurements. With p at 4 ms, each *S* was able to distinguish two separate lines of light at 15 degrees nasal while fixating the 5-mm luminous circle. As p was increased, the two lines became blurred into a single vaguely defined area, and the ascending two-point threshold (TP_{Asc}) was recorded when *S* reported he was no longer able to distinguish two separate lines. As p was further increased, a left-to-right apparent movement of, and within, the vaguely defined area could be observed, and the peripheral apparent movement threshold (M_P) was recorded when *S* first reported such a movement. Then p was decreased until *S* again re-

ported two separate lines. This was taken as the descending two-point threshold (TP_{Desc}).

The instructions, aimed at minimizing differences between ascending and descending thresholds, follow:

For A_{Asc}: Now you see two lines of light going on and off at exactly the same time, exactly together. Let me know as soon as you notice that they are not going on exactly together any more, let me know as soon as you notice that they are out of order even the slightest bit.

For M_{Asc}: Now you see two lines out of order, sort of chasing each other. Let me know as soon as you cannot see two lines any more. Let me know as soon as it looks just like one line that's moving across from left to right.

For M_{Desc}: Now you see only one line that is moving from left to right. Let me know as soon as you can tell there are two lines again, as soon as you can tell it is not just one line that is moving from left to right.

For A_{Desc}: Now you see two lines blinking out of order. Let me know as soon as you notice that they are going on and off exactly together, that they are not out of order even the slightest bit.

For TP_{Asc}: Keep looking at the blue light (the luminous fixation point). Let me know as soon as you cannot make out two separate lines any more, let me know as soon as it looks just like one blotch of light going on and off. Don't take your eyes off the blue light.

For M_P: Keep on looking at the blue light. Let me know as soon as the blotch looks like it is moving a little from left to right, as soon as it doesn't just seem to go on and off, as soon as it seems to move across a little bit. Don't take your eyes off the blue light.

For TP_{Desc}: Keep on looking at the blue light. Let me know as soon as it no longer looks like just a blotch, but as soon as you can make out two separate lines, or two distinct lights again. Don't take your eyes off the blue light.

The estimated threshold for each *S* in this experiment is always the mean of six threshold measurements. The *V* score is always taken as the logarithm of the variance of the distribution [2] of measurements around the mean, i.e., around the threshold. Nonparametric rank methods (14) were indicated in view of the sometimes erratic estimates of spread (*V* scores) of the psychometric functions. The .05 level of significance was adopted as the criterion for rejecting the null hypothesis of no difference between groups.

[2] A useful property of such a distribution of log variance scores is that their variance will depend *only* on the degrees of freedom in the estimate of the *V* score (1). If the psychometric function is a cumulative normal curve, then the variance of $\log_{10}s^2$ with 5 *df* is .0938 (if the curve is not a grossly non-normal cumulative curve, this is still approximately true). This makes it possible to compare the variance obtained for any of the average *V* scores in Table 2 with the known constant variance of .0938 by the *F* ratio and establish the significance level.

Results and Discussion

Higher average thresholds were found in the brain-injured than in the control group (Table 1). For ascending asynchronism thresholds, it required an average pause of 44.5 ms for the brain-injured to notice that the two lights were not presented simultaneously, while an average pause of only 34.0 ms between the onset of the two lights was noticed by the control group. The difference of 10.5 ms between the groups is significant at the .05 level and supports the hypothesis that visual efficiency is impaired in the brain-injured. The average pause between the two lights for the ascending apparent movement threshold was 110.8 ms for the brain-injured group and 83.9 ms for the controls; the difference of 26.9 ms between the two groups is significant at the .01 level.

Table 1

Average Foveal Asynchronism Thresholds (A_{Asc} and A_{Desc}), Foveal Apparent Movement Thresholds (M_{Asc} and M_{Desc}), Peripheral Two-Point Thresholds (TP_{Asc} and TP_{Desc}), Peripheral Movement Thresholds (M_P), and Foveal-Minus-Peripheral Movement Scores ($M_{Asc} - M_P$) in Milliseconds for the Brain-Injured and Control Groups of Ten *S*s Each

	Brain-injured		Control		
Thresholds	$\bar{X}$	*SD*	$\bar{X}$	*SD*	Diff.
Foveal					
A_{Asc}	44.5	11.2	34.0	8.1	10.5*
A_{Desc}	33.3	9.6	30.0	4.9	3.3
M_{Asc}	110.8	20.0	83.9	20.7	26.9**
M_{Desc}	81.2	10.9	72.1	15.6	9.1
Peripheral					
TP_{Asc}	53.7	17.5	52.8	19.5	.9
TP_{Desc}	53.6	10.6	41.6	19.3	8.0
M_P	92.7	22.0	97.3	32.0	4.6
Foveal-minus-peripheral					
$M_{Asc} - M_P$	18.1	17.8	−13.4	28.0	31.5**

* Significant at the .05 level of confidence.
** Significant at the .01 level of confidence.

Table 2

Average *V* Scores (Log Variance Scores) for Foveal Asynchronism (A_{Asc} and A_{Desc}), Foveal Apparent Movement (M_{Asc} and M_{Desc}), Peripheral Two-Point (TP_{Asc} and TP_{Desc}), and Peripheral Apparent Movement Measurements for the Brain-Injured and Control Groups of Ten *Ss* Each

	Brain-injured		Control	
Thresholds	$\bar{X}$	F^a	$\bar{X}$	F^a
Foveal				
A_{Asc}	2.05	2.62	1.69	2.10
A_{Desc}	2.19**	.42	1.74	1.44
M_{Asc}	2.16	5.03	2.11	1.96
M_{Desc}	2.02	1.55	2.06	2.11
Peripheral				
TP_{Asc}	2.44**	.78	1.98	1.67
TP_{Desc}	2.49	1.64	2.21	.69
M_P	2.60	1.26	2.23	1.78

** Significant at the .01 level of confidence.
[a] If in fact all *Ss* within a group had the same intra-individual variability for a particular threshold, then $F_{(9,\infty)} \geq 1.88$ would occur only 1 time in 20.

When each *S*'s peripheral movement threshold (M_P) is compared with (i.e., subtracted from) his own corresponding foveal movement threshold (M_{Asc}), the brain-injured decreased their thresholds by an average of 18.1 ms, while the controls increased their thresholds by 13.4 ms. The difference between the two groups in this respect is significant at the .01 level and suggests that just as in previous CFF experiments, changes in thresholds due to altered conditions may yield good discriminators between the two groups.

Six out of seven average *V* scores of the brain-injured group were larger than those of the control group (Table 2). The *V* scores for foveal descending asynchronism thresholds and peripheral ascending two-point thresholds discriminated between the two groups at the .01 level of significance. Such findings throughout these studies support one of the major hypotheses that the brain-injured show increased intra-individual variability.

The visual thresholds in this experiment on the whole did not discriminate as sharply between brain-injured and control groups as did the light perception thresholds of the previous investigation (7). In view of the fact that all four *V* scores associated with the light perception thresholds discriminated between the two groups at the .01 level, the findings here serve to strengthen the major hypotheses but hold no additional promise as potential diagnostic discriminators. They may be of some theoretical interest however. Differences in apparent movement thresholds in the previously reported experiment, where the exposure time of the left light was the only experimental variable, are readily attributable to difference in exposure time required for light perception thresholds between the two groups. In this study where the pause between the two lights was the only experimental variable altered, central mechanisms may well have contributed to differences between the groups. Klein and Krech (4) and Eysenck (3) suggest that distinct nosological groups have characteristic differences in central after-effects. Experiments designed to test specific figural after-effect hypotheses related to magnitude, speed of arousal, and persistence of after-effects are reported elsewhere (7).

Summary

Differences in asynchronism, apparent movement, and peripheral two-point thresholds found between brain-injured (pyramidal tract damaged) and control children support the two major hypotheses of impairment in visual efficiency, and increased intra-individual variability in the brain-injured. The possible significance of the apparent movement threshold findings is discussed.

Received July 15, 1957.

References

1. Bartlett, M. S., & Kendall, D. G. The statistical analysis of variance-heterogeneity and the logarithmic transformation. *J. Roy. statist. Soc.* (*Suppl.*), 1946, **8**, 128–138.
2. Birch, H. G. Personal communication. 1955.
3. Eysenck, H. J. Cortical inhibition, figural after-effect, and theory of personality. *J. abnorm. soc. Psychol.*, 1955, **51**, 94–106.
4. Klein, G. S., & Krech, D. Cortical conductivity in the brain-injured. *J. Pers.*, 1952, **21**, 118–148.
5. Lilienfeld, A. M., & Pasamanick, B. Association of maternal and fetal factors with the development of epilepsy: I. Abnormalities in the prenatal and paranatal periods. *J. Amer. med. Ass.*, 1954, **155**, 719–724.

6. Lilienfeld, A. M., & Pasamanick, B. The association of prenatal and paranatal factors with the development of cerebral palsy and epilepsy. *Amer. J. Obstet. Gynecol.*, 1955, **70**, 93–101.
7. Mark, H. J. Studies of perception in brain-injured children. Unpublished doctoral dissertation, Johns Hopkins University, School of Hygiene and Public Health, 1955.
8. Mark, H. J., Meier, P., & Pasamanick, B. Variability of critical flicker fusion thresholds in brain-injured children. *A.M.A. Arch. Neurol. Psychiat.*, in press.
9. Morgan, C. T., & Stellar, E. *Physiological psychology*. New York: McGraw-Hill, 1950.
10. Pasamanick, B., & Lilienfeld, A. M. Maternal and fetal factors in the development of epilepsy: II. Relationship to some clinical features of epilepsy. *Neurology*, 1955, **5**, 77–83.
11. Pasamanick, B., & Lilienfeld, A. M. The association of maternal and fetal factors with the development of mental deficiency: I. Abnormalities in the prenatal and paranatal periods. *J. Amer. med. Ass.*, 1955, **159**, 155–160.
12. Rogers, M. E., Pasamanick, B., & Lilienfeld, A. M. Prenatal and paranatal factors in the development of childhood behavior disorders. *Acta psychiat., Kbh.*, 1955, Suppl., **102**, 157.
13. Roush, R. G., & Urbanski, E. T. Universal medical timer and pulse stimulator. *Electronics*, 1953, **26** (II), 154–157.
14. Wilcoxon, F. *Some rapid approximate statistical procedures*. New York: American Cyanamid, 1949.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Opinions of Psychologists and Psychiatrists About Mental Health Problems[1]

Jum Nunnally
University of Illinois

The research reported here grew out of a larger program concerning the communication of information about mental health to the general public.[2] The first step in this larger program was to learn the state of public understanding about mental health problems. In order to do this, it was necessary to learn what the public thinks about mental health problems and to compare this with the opinions of experts in the mental health field. The studies of public opinion are reported elsewhere (1). This paper presents results obtained from a study of experts.

Questionnaire Construction

The term "opinions" refers in the specific sense to segments of verifiable knowledge. For example, the statement, "There are more men than women in mental hospitals," is classified as an opinion, because it could be verified one way or the other. On the other hand, a statement such as, "I am afraid to be around anyone who has had a nervous breakdown," indicates a personal reaction, or attitude. The studies of attitude toward mental health problems came at a later stage. The present study is largely concerned with opinion.

A measuring instrument was needed which was general enough to be used not only with diverse segments of the general population, but also with mental health experts. For this purpose, it seemed most appropriate that the questionnaire items should be selected from the opinions and beliefs currently held by the persons in our two sources. Opinion statements were sampled from a wide collection of materials including personal interview data, recorded interviews with the mothers of children in psychotherapy, and interviews with the relatives of mental hospital patients.[3] Additional items were sampled from mental hygiene text books and other professional publications. Samples of content from the mass media—radio, television, newspapers, and magazines—were searched for their mental health content. For example, when the "psychiatrist" in one of the radio "soap operas" advised his patient to take a long vacation, this was written down as one opinion about ways in which to treat mental health problems. In this way over 3,000 statements were gathered relating to the causes, treatment, symptoms, prognosis, incidence, and social significance of mental health problems.

The opinion statements which were collected range from the belief that "insanity" results from over-eating to the belief that psychiatrists are older than other physicians. Only those opinion statements were recorded which were expressed in a language the average person could understand. Consequently, the sample of opinions largely contained the simpler and more abbreviated opinions. Thus, there may be many opinions of a complex nature held by mental health experts which have

[1] The research was supported by a special grant from the National Institute of Mental Health, U. S. Public Health Service.

[2] J. C. Nunnally and C. E. Osgood are co-directors of the project. Members of the project staff during the time of the research were G. J. Suci, W. L. Taylor, and R. J. Wolff. The research assistants, T. R. Husek and Carol Tucker, gave very helpful support to the work.

[3] For lending us materials for the sampling of items, we would like to thank, among others, Angus Campbell, John Clausen, Elliot Freidson, and Shirley Starr.

no counterparts in our sample of opinion statements.

The larger collection of opinions was reduced to a group of 240 by removing apparent duplicates in meaning. The 240 statements were then cast in questionnaire form. A seven-point rating scale was provided for each, ranging from "agree" to "disagree." Several examples of the opinion statements are:

Psychiatrists try to teach patients to hold in their strong emotions.

Good emotional habits can be taught to children in school as easily as spelling can.

Adult problems are less important in causing emotional disorders than the individual's childhood experiences.

The questionnaire was administered to a sample of the general population in the Champaign-Urbana area. It was found that almost everyone either agreed completely or disagreed completely with 60 of the items. These were withheld from further analysis. The remaining 180 items were subjected to factor analysis. The 11 factors obtained in this way were used to describe popular beliefs about mental health problems. On the basis of the factors, a shortened version of the questionnaire was constructed containing 50 items. The 50-item form was then administered to samples of the population in Eugene, Oregon, and in Knoxville, Tennessee. This form was also used to study the opinions of experts.

The Sample of Experts

The sample consisted of 300 psychologists and psychiatrists in the United States. The 150 psychologists were randomly drawn from the diplomates in the American Psychological Association, with 75 of these being taken from the clinical section and the remaining 75 from the counseling and guidance section. The psychiatric sample consisted of 150 members of the Group for the Advancement of Psychiatry. It was thought that a sample obtained in this way could be said to represent at least a good portion of the expertness in psychology and psychiatry. Each person was mailed a copy of the questionnaire with a slightly different instruction from that used in studies of the lay population. Instead of asking each of the experts whether he agreed or disagreed with the opinion statements, he was asked whether each opinion was worthy of being advocated to the public, or used in programs of public information. Consequently, the experts used a seven-point rating scale which ranged from "support" to "repudiate."

It was explained to the experts that the middle of the continuum would be used to indicate that a particular idea should be ignored in public information efforts. Essentially, the experts were asked what messages should be conveyed to the public. In the strictest sense, the results of the study do not show what the experts believe but, more precisely, what they think the public should be told. A subsequent study of a smaller group of experts using the original "agree-disagree" questionnaire form showed such a strong correspondence with the results obtained from the "repudiate-support" form that it is reasonable to assume that the experts were responding mainly in terms of what they personally believed.

Returns from the sample were good. Of the 300 questionnaires mailed out, 201 were returned, 176 in time for our analysis. The group which was analyzed contained the responses of 86 psychologists and 90 psychiatrists.

Analysis of Data

The principal reason for getting the expert sample was to gauge the public understanding and to locate the groups in the general population whose beliefs appeared to be the most anachronistic in comparison to expert opinion. These and other findings are reported separately (1). The concern here is within the expert sample, to find the points on which they agree and disagree. First, the mean position of each of the three subgroups (two groups of psychologists and the psychiatrist group) of the expert sample on each of the questionnaire items was found. Next, standard deviations for the individual items were determined over the whole expert sample. Finally, a factor analysis was performed on 22 of the 50 questionnaire items. The particular 22 items were chosen because they showed the largest standard deviations, or disagreement, among experts.

Table 1

Rotated Orthogonal Factors

Item	Factor I	II	III	IV	V
1	.24	.29	.00	−.15	.03
2	.01	.14	.35	−.23	−.18
3	.10	.27	−.02	.14	.12
4	−.22	.48	.03	−.10	−.17
5	−.03	.00	−.09	.35	−.20
6	.18	.37	.00	.22	.05
7	.57	−.03	.16	−.04	−.05
8	.40	.08	.07	.00	.43
9	.60	.03	−.12	.03	.05
10	.02	.34	−.10	−.14	.25
11	.10	.20	.08	.39	.02
12	−.10	−.12	.30	.26	.01
13	.05	.30	.17	−.09	.04
14	.22	−.23	.21	.25	.07
15	−.07	.08	.35	−.17	.29
16	.04	−.28	.10	.38	.01
17	−.23	.35	.20	.09	−.01
18	.43	−.22	.00	.10	−.10
19	.01	.33	.18	.04	−.06
20	.31	−.04	−.03	−.35	.22
21	.42	−.20	−.11	.07	−.37
22	−.19	.46	−.05	−.02	.00
23	.03	.23	.42	−.01	.00

Results

The worst fear was that the experts would show themselves no more in agreement on the questionnaire items than members of the general public. This proved not to be the case. The item variances of the experts were in all cases lower than those of the lay public and generally tended to be less than one-half [4] as large. Another way of looking at the amount of agreement among experts is in terms of the percentage of persons who disagreed with the majority decision on each item. If more than 50% of the respondents indicate that a particular item should be advocated to the public, then the percentage of persons who mark one of the "repudiate" categories, positions 1, 2, and 3, on the scale, indicates the percentage of disagreement with the majority decision. Whereas on half of the 50 items the experts show less than 10% disagreement with the majority decision, the general population shows this degree of agreement on none of the items.

The next concern was whether systematic differences would be found between psychologists and psychiatrists, or between the two groups of psychologists. The average responses (the arithmetic means) by psychologists and psychiatrists proved to be very much alike. Some small differences between professions will be shown in respect to the factor analysis work. Likewise, the average responses by the clinical and counseling diplomates were very much alike, with none of the differences approaching statistical significance.

The third consideration was whether there was sufficient disagreement on some of the items, in the context of generally high agreement, to point to schools of thought about mental health. For this purpose, the 22 items which showed the largest variance were intercorrelated and factor analyzed [5] (see Table 1). The product-moment correlation formula was used for all coefficients.

As a twenty-third variable in the correlation matrix and in the factor analysis, the profession of each subject was included. A score of 1 indicated that the subject was a psychiatrist and a score of 2 indicated that the subject was a psychologist.[6] The professional status variable was then factored in with the 22 items. The loadings of the twenty-

[4] Variances were averaged over the 50 items for experts and the lay population separately. The average variance for the lay population was found to be 2.3 times as large as that for the experts.

[5] The correlation matrix was factor analyzed by the principal axes method (cf. 2), sometimes called the principal components method. Communality estimates were used to fill the diagonal elements of the matrix. These were estimated by selecting the highest correlation coefficient, regardless of sign, in each column of the matrix. The five factors with the largest variance were placed through a series of single-plane rotations to achieve the orthogonal solution in Table 1. Copies of the correlation matrix, the unrotated factor matrix, and the questionnaire will be sent on request. Requests should be directed to J. C. Nunnally, Department of Psychology, University of Illinois, Urbana, Illinois. Or the material can be obtained by ordering Document No. 5486 from the American Document Institute, remitting $1.25 for microfilm or $1.25 for photocopies.

[6] When the product-moment correlation formula is applied to the relationship between one discrete and one continuous variable, the obtained coefficient is the same as the point biserial coefficient.

third variable indicate the extent to which one profession is more inclined to believe a factor than is the other profession. A positive loading means that the factor is held higher by psychologists than by psychiatrists. A negative loading means that the particular factor is held higher by psychiatrists.

The most highly loaded items on the first factor are:

Factor I

Loading		*Item*
.60	[9]	Psychiatrists try to show the mental patient where his ideas are incorrect.
.57	[7]	The main job of the psychiatrist is to explain to the patient the origin of his troubles.
.43	[18]	The good psychiatrist acts like a father to his patients.
.42	[21]	Psychiatrists try to teach mental patients to hold in their strong emotions.
.40	[8]	Psychiatrists have to have a good sense of humor in order to help their patients.

Factor I concerns the psychiatrist's[7] role, and, probably, the role of psychotherapists generally. More particularly, the factor appears to concern the degree of "directiveness" that is thought to characterize the effective therapist. The items picture the effective therapists as teaching, correcting, and explaining. The factor illustrates the well-known professional controversy between those who advocate a more directive, interventionist role for the therapist, as opposed to those who advocate a role of listening and understanding. It can be seen from Table 1 that Factor I is not more characteristic of psychiatrists than psychologists. Variable 23, professional affiliation, has a loading of only .03 on the first factor. Instead, Factor I demonstrates that there is a division of opinion within both professions concerning the extent to which the therapist should direct and control the therapy situation.

Factor II

Loading		*Item*
.48	[4]	Physical rest will not prevent a mental disorder.
.46	[22]	A change of climate seldom helps an emotional disorder.
.37	[6]	Adult problems are less important in causing emotional disorders than the individual's childhood experience.
.35	[17]	Helping the mentally ill person with his financial and social problems will not cure his disorder.

The content of Factor II concerns the causes of mental disorders and the influences that make for better or worse emotional adjustment. Persons who agree strongly with these items would seem to favor an explanation of mental disorder in terms of the individual's childhood and his history of personal experience. A person who held the opposite point of view, one who would disagree with the items above, would look for the causal factors in the pressures in the individual's immediate environment. The extreme of this point of view would be that the individual's emotional adjustment is a function only of contemporary external influences. The extreme opposite point of view would be that the immediate environmental pressures are of no importance in explaining the individual's status of adjustment.

Variable 23, professional affiliation, has a small loading of .23 on Factor II. If this loading is not due to chance,[8] there is a slight tendency for psychologists to endorse Factor II more than psychiatrists. However, even if the difference between professions is real, the difference within professions is the far larger consideration.

[7] The term "psychiatrist" was used throughout the questionnaire to refer to an expert in the mental health field. We would have liked to use different titles, like psychiatrist, psychologist, and psychoanalyst, but this would have increased the items beyond the number that could be used in survey work. In scanning numerous personal interviews, we found that the psychiatrist is the person most often mentioned by members of the public as a person who deals with mental health problems. Also, we found that the public makes little distinction between the subprofessions in the mental health field.

[8] There are no exact error formulas that can be applied to rotated factor loadings. An analogous error term can be obtained from the standard error of the correlation coefficient. With an N of 176 in this case, the standard error would be .076, with a coefficient of .19 being significant at the .01 level. By this very approximate standard, the loading of .23 would be indicative of a real difference between the two professions.

Factor III

Loading		*Item*
.42	[23]	Professional affiliation (variable 23 in Table 1).
.35	[15]	A poor diet does not lead to feeblemindedness.
.35	[2]	Physical exhaustion does not lead to a nervous breakdown.
.30	[12]	Most of the time psychiatrists have difficulty in telling whether or not a patient's mental disorder is curable.

Factor III contains most of the common variance for the variable of professional affiliation. Rather than being a content factor, in which there should be some centralness of meaning, the factor contains an array of the items that psychologists tend to endorse more strongly than do psychiatrists. As mentioned previously, there are few items that differentiate the professions; and the differences, as evidenced in the loadings above, are small. The factor shows that psychologists are more skeptical of organic and "physical" explanations of mental disorder; i.e., diet and physical exhaustion, than are psychiatrists. The last listed item in Factor III, that concerning the psychiatrist's difficulties in prognosis, perhaps reflects more enthusiasm by some psychiatrists about their clinical prowess than would be granted by some psychologists.

Factor IV

Loading		*Item*
.39	[11]	Mental illness is one of the most damaging disorders that a person can have.
.38	[16]	Most social workers go into their field because of their own misfortunes in life.
.35	[5]	The brains of the feebleminded are smaller than those of normals.
— .35	[20]	When a person is recovering from a mental illness it is best not to discuss the treatment that he has had.

No sensible interpretation could be found for Factor IV. It is submitted to the reader for his inspection.

The loadings on Factor V are too small to be considered without additional evidence for the statistical stability of the factor. Of the common variance explained by the five factors, the percentages of variance for each were, respectively, 29.8, 28.0, 13.7, 16.6, and 12.1.

The findings in this research are restricted to the area of opinions being studied. A sample of opinions drawn more directly from what experts think and say might explain more about divisions of opinion in the mental health field. The most that can be said at this time is that there are some differences of opinions along the several lines which the factors have indicated. Other differences might be observable if new items were studied and additional factors were found.

Summary

The article concerns a questionnaire study of the opinions of psychologists and psychiatrists about mental health problems. The two professions were in agreement about most of the issues. A factor analysis was made of the combined responses for the two professions. Two of the obtained factors demonstrate schools of thought about mental health problems. A third factor contains a number of items which serve to differentiate the opinions of psychologists and psychiatrists.

Received June 3, 1957.

References

1. Nunnally, Jum. The communication of mental health information: A comparison of the opinions of experts and the public with mass media presentations. *Behav. Sci.*, 1957, **2**, 222–230.
2. Thurstone, L. L. *Multiple factor analysis.* Chicago: Univer. of Chicago Press, 1947.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Some Social-Psychological Characteristics of a Child Guidance Clinic Caseload[1]

Jack L. Roach[2]
Guidance Center of Buffalo

Orville Gurrslin
Youth Board of Buffalo

and Raymond G. Hunt
University of Buffalo

A number of recent studies have indicated that social-psychological variables play a significant role in determining the utilization of mental health facilities by members of the community (1, 7, 9, 10). While such a phenomenon is generally in accord with certain theoretical expectations (4), it is difficult to draw generalized conclusions from the available literature. By and large, studies in this area have dealt with tendencies in specific geographic areas and, usually, specific mental health facilities. As a result, the data obtained may be seriously biased as a function of idiosyncratic features of the particular area or agency involved. Thus the degree to which sources of data are representative of the universe to which this data becomes generalized is an especially critical issue in studies of this type.

The sampling problem can perhaps be resolved in a number of ways, probably the most efficient of which is replication of such studies with independent samples in different geographical areas. The pooled findings of these studies then constitute descriptive steps permitting analytic or systematic generalization concerning the sociology and social psychology of mental health and institutional structures in that area.

[1] The authors wish to thank A. V. Cutter, director, and the staff of the Guidance Center of Buffalo whose cooperation made this study possible.

[2] Now with the New York State Department of Social Welfare.

The present investigation then, while not a direct replication of any particular study, is an attempt to investigate the relationship between a number of social-psychological characteristics found to be of significance by other investigators and the caseload of a community child guidance clinic.

Method

Case record information was obtained from 246 cases served in some capacity at the Guidance Center of Buffalo during 1954 and 1955. The Guidance Center is a community supported child guidance clinic with services available, through an open referral system, to the residents of Erie County, New York. Fees are adjusted according to ability to pay, there being no reason for a family to be refused service because of low income.

The study includes the entire child caseload active at some time during the period covered with the exception of cases wherein the child was not residing with his own parent(s), either natural or adoptive. Thus, institutionalized or foster-home children were excluded from the study.

Data pertaining to several social-psychological characteristics of the families receiving clinic services were extracted primarily from the application form filled out by the parents at intake. As generally happens in studies of this sort, case records do not always yield complete information about every case. As a

result, in each descriptive category employed there was a varying number of cases which could not be ordered exactly. After careful consideration, however, the authors deem it unlikely that the findings reported below would be seriously altered were all of the unknown cases known. Comparative census data, where appropriate, were drawn from the Statistical Abstract of the United States (11), and from records of the Council of Social Agencies of Buffalo and Erie County.

Seven characteristics presumed to be of significance were selected for analysis. These were: (*a*) sex of child, (*b*) family size, (*c*) occupation of major wage earner, (*d*) area of residence, (*e*) race, (*f*) religious affiliation, and (*g*) symptom picture.

The purpose of the present study was entirely descriptive (empirical). No attempt was made, at this time, to establish relationships between any of the variables, though a number of interrelations are likely as suggested below.

Results

It was found that of the 246 cases included in this study, 72% were male. This finding is consistent with data reported by several other investigators (1, 7), thus indicating a tendency toward disproportionate representation of males in child guidance clinic patient populations.

Table 1

Occupational Distribution of Major Wage Earner in Clinic Families

(N = 241)

Occupational group	Source (in per cent)	
	Clinic[a]	Census[b]
Professional and technical	21	11
Proprietors, managers, officials	20	12
White collar, clerical	19	24
Skilled workers	19	15
Semiskilled workers	11	31
Unskilled workers	10	7
Total	100	100

Note.—$\chi^2\% = 24.77$, $df = 5$, $p = <.001$.
[a] Unknowns (7% of total) were distributed among categories according to observed distribution of known cases.
[b] Census figures refer to New York State (1950).

Table 2

Quartile Ratings of Desirability of Residential Area

(N = 196[a])

Quartile ranks	Source (in per cent)	
	Observed (clinic)	Expected
I	52	25
II	11	25
III	24	25
IV	13	25
Total	100	100

Note.—$\chi^2\% = 57.82$; $df = 3$; $p = <.001$.
[a] The bulk of unknown cases (20%) omitted from the analysis came from tract areas not rated but primarily desirable suburban areas.

A similar lack of representativeness was observed with respect to the racial characteristics of the clinic population—only 1% of which was nonwhite. In both instances, the significance (statistical) of the data is obvious.

Table 1 reports the results of a comparison between the occupational distribution of the major wage earner in the clinic population and the distribution of such occupations in the general population. The classification employed was a modification of occupational classifications frequently reported in the sociological literature (2). Farmers and unemployed workers were excluded from the analysis owing to the very low frequency of each in the clinic population (two and three cases, respectively). Moreover, it would be expected that few farmers would be represented owing to the urban character of the area served by the clinic. Hence, to avoid the introduction of spurious differences, these groups were omitted from analysis.

Table 1 indicates that the agency is serving all occupational groups within the general population, but with a considerable overrepresentation of the "upper" occupational categories.

The data reported in Table 2 were obtained by rating the area of residence of each clinic family along a dimension of desirability, according to a quartile scale developed by the Council of Social Agencies of Buffalo and Erie County (3). This scale is based upon a

Table 3

Religious Affiliation of Clinic Families
($N = 246$)

Religion	Source (in per cent)	
	Clinic[a]	Census[b]
Protestant	57	40
Catholic	32	57
Jewish	11	3
Total	100	100

Note.—$\chi^2_{.01} = 33.50$; $df = 2$; $p = <.001$.
[a] Unknowns (20%) were distributed according to observed frequencies of known cases.
[b] Census figures refer to metropolitan Buffalo (1955).

variety of objective characteristics of census tracts in Buffalo and vicinity (e.g., percentage of dilapidated dwellings). It should be noted that a rating of I corresponds to "most desirable" and a rating of IV to "least desirable."

It can be seen from Table 2 that a heavy proportion of clinic families come from the most desirable residential areas. It should be borne in mind that the variables summarized in Tables 1 and 2 involve dimensions widely used as indices of social-class position.

Data pertaining to the religious composition of the clinic caseload are presented in Table 3. Both the present study and Anderson and Dean (1) agree generally in finding that Protestants comprise the largest single clinic group. Both this and the Jewish group are overrepresented in the present clinic population, while Catholics are considerably underrepresented.

In Table 4 are presented data relating to the number of children in clinic families. It is clear that the obtained difference is accounted for mainly by the underrepresentation of single child families in the clinic population. It is of considerable interest that only 15% of the case population is made up of "only" children—a finding contrary perhaps to expectancies derived from much child development literature (5).

Table 5 represents a classification of the symptomatic picture characterizing the child at intake. The classification was developed by the authors. Cases were ordered to categories on the basis of dominant symptom patterns. When no particular type of symptom was dominant, the case was classified as "mixed," hence this category is somewhat more heterogeneous than the others. As far as possible, ratings were based upon descriptions of the child's behavior provided by the parent(s). An effort was made to sift out interpretive statements originating with clinic personnel. It was hoped thereby to lend greater objectivity to the data.

Evident in Table 5 is the marked incidence of children with aggressive behavior as the predominant symptom. This seems generally consistent with the findings of Stevens (10) and of Anderson and Dean (1). Thus, aggressive and acting-out behavior apparently com-

Table 4

Number of Children in Clinic Families
($N = 246$)

Number of children	Source (in per cent)	
	Clinic[a]	Census[b]
1	15	35
2	41	33
3	20	18
4 and more	24	14
Total	100	100

Note.—$\chi^2_{.01} = 20.86$; $df = 3$; $p = <.001$.
[a] Unknowns (3%) were distributed according to observed frequencies of known cases.
[b] Census figures refer to New York State (1955).

Table 5

Classification of Dominant Symptomatic Behavior at Intake
($N = 246$)

Dominant symptoms	*f*	%
Aggressive	86	35
Nonaggressive[a]	69	28
Somatic complaints[b]	23	9
Mentally retarded[c]	35	14
Mixed	25	10
Unknown	8	4
Total	246	100

[a] For example, withdrawal, daydreaming, "silly behavior," excluding somatic functions when dominant complaint.
[b] For example, asthma, pains in stomach, when dominant complaint.
[c] Cases classified here when dominant complaint was retardation, suspicion thereof, or poor academic performance.

prises the largest single basis for clinic referral, perhaps because such behavior is most disturbing to various cultural agents.

Finally, it is perhaps worth noting that 25% of the patient population included in this study had had some kind of prior social agency contact as reflected by registration with the Social Service Exchange.

Discussion and Conclusions

A general conclusion which emerges from the findings reported here is that the clinic from which the data was drawn is not serving equally all segments of the population. It is clear that the clientele served represents, primarily, a white, Protestant group of middle to high socioeconomic status. Such also is the conclusion drawn by other similar studies.

Further, it is quite conceivable that most of the observed instances of unrepresentativeness of clinic populations are a direct function of the middle-class character of such populations. If, for example, the population served is largely middle class, it is not at all surprising that nonwhites will form a very small percentage of clinic clientele. Likewise, the convergence between Protestant and general middle-class values pointed out by many writers (Davis, 4; Merton, 8; Williams, 12), forms one plausible explanation for the observed distribution of religious groups in the clinic population.

Further, the fact that a large percentage of Catholics is included in the lower social classes renders understandable the underrepresentation of this group in the population of urban child guidance clinics, although other factors might also be operative with regard to this group. The same considerations would hold, in reverse, as an explanation for the observed overrepresentation of Jews.

Similarly, this "middle-class phenomenon" might be of considerable significance with respect to the preponderance of male children served (Green, 6), and also with regard to the dominant type of symptom patterns observed. Regarding this latter, it seems likely that instances of acting-out behavior will be a source of greater concern among middle-class than among working-class families.

It thus becomes apparent from a growing body of literature similar to the present study, that cultural mental health facilities are being utilized primarily by middle- and upper-class members of society. The independent variables determining this phenomenon are as yet, however, not clearly understood and obviously call for empirical elucidation.

Received June 24, 1957.

References

1. Anderson, F. N., & Dean, H. C. Some aspects of child guidance clinic intake policy and practice. U. S. Dept. of Hlth, Educ. & Welf., *Publ. Hlth Monogr.*, 1956, No. 42.
2. Bendix, R., & Lipset, S. *Class, status and power.* Glencoe, Ill.: Free Press, 1953.
3. Council of Social Agencies of Buffalo and Erie County, N. Y. *Buffalo self study of recreation, informal education and group work services.* Buffalo, N. Y.: Author, 1956.
4. Davis, K. Mental hygiene and the class structure. In P. Mullahy (Ed.), *A study of interpersonal relations.* New York: Hermitage House, 1949. Pp. 364–386.
5. Epperson, J. A. Some basic principles of direct work with children. *Child Welfare,* 1952, **21**, 3–8.
6. Green, A. The middle-class male child and neurosis. In L. Wilson & W. Kolb (Eds.), *Sociological analysis.* New York: Harcourt, Brace, 1949. Pp. 236–248.
7. Hollingshead, A., & Redlich, F. Social stratification and psychiatric disorders. *Amer. sociol. Rev.,* 1953, **18**, 163–169.
8. Merton, R. K. *Social theory and social structure.* Glencoe, Ill.: Free Press, 1949.
9. Michigan Department of Mental Health. Differential utilization of the facilities of a Michigan child guidance clinic. *Mich. Dept. ment. Hlth res. Rep.,* Mar. 1955, No. 17.
10. Stevens, S. An ecological study of a child guidance clinic intake. *Smith Coll. Stud. soc. Wk.,* 1954, **25**, 73–84.
11. U. S. Dept. of Commerce, Bureau of the Census. *Statistical abstract of the United States.* Washington: U. S. Govt. Printing Office, 1956.
12. Williams, R. *American society.* New York: Knopf, 1951.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

The Weighted *Q* Sort: A Procedure for Quantitatively Estimating Emotional Disturbance and Personality Change[1]

Marvin Reznikoff and Laura C. Toomey
Institute of Living, Hartford, Connecticut

While the usefulness of the *Q* sort has been clearly demonstrated in expressing individual personality profiles in quantitative terms, the technique is not equally effective for quantitative representations of personality change or degree of emotional maladjustment. In a typical application of this forced choice rating method, *Q*-sort personality evaluations are done on patients before and after therapy and the sorts compared. Coefficients of correlation indicate if personality changes, per se, have occurred, but do not reflect quantitatively the direction and magnitude of such changes; neither do *Q* sorts used in this manner provide direct estimates of levels of adjustment which can be contrasted. Only Dymond (3) has published a method for quantitatively evaluating adjustment using an "external criterion." Dichotomizing her *Q*-array statements into "adjustment" and "maladjustment" categories, she arrives at an adjustment score by counting the adjustment items characteristic of the patient and the maladjustment items not characteristic of him.

As one aspect of a large project involving the evaluation of psychological changes associated with psychiatric treatment, a modification of the *Q* sort was developed which yields, along with the usual personality picture, a single composite score revealing degree of emotional disturbance. Additionally, this "adjustment index" readily lends itself to an over-all quantitative determination of personality change. In contrast to Dymond's method of broadly classifying items as indicative of either adjustment or maladjustment, this procedure places items at appropriate points on a graduated adjustment-maladjustment scale. The modified technique designated as "the weighted *Q* sort" is described in this paper.

[1] This investigation was supported by the Office of Naval Research under Contract No. NONR (00) (01).

Development of the Method

As a first step, a group of clinical formulations were selected from Beck's psychodynamic outline of schizophrenia (1) and modified to some extent. Since Beck's conceptual schema is based solely on Rorschach personality correlates, pertinent clinical formulations were also included which could be determined more readily from other projective tests, as well as from information elicited during therapeutic sessions. Finally, some statements characteristic of personality in general were added, in order to more fully represent the normative end of the continuum, and provide a schema applicable to nonschizophrenic as well as schizophrenic subjects. All 44 clinical statements were then typed on individual 3" × 5" cards and sorted into a seven-category forced normal distribution by three experienced clinicians. These judges, working independently, rated these items on a continuum from most indicative of severe emotional disturbance (Category 7) to least indicative of severe emotional disturbance (Category 1). The specific sorting instructions provided were:

> Please sort these 44 clinical formulations into seven categories. Category 7 is to represent those clinical formulations most indicative of a serious psychological disturbance. The clinical formulations in the remaining six categories are to be thought of as progressively less indicative of a serious mental disturb-

ance with those items in Category 1 considered to be least pathological in their implication. The number of cards to be placed in each of the categories is as follows:

Category	7	6	5	4	3	2	1
Number of cards	3	6	8	10	8	6	3

The correlations between the ratings of the three clinicians taken two at a time were .84, .75, and .73, indicating a statistically significant degree of agreement ($p < .001$) in evaluating the relative pathological implications of the clinical formulations. Using the totals of the three independent ratings for each of the 44 items distributed in an approximate normal curve, a weight or "pathology score" was assigned to every clinical formulation. The three clinical formulations having the highest total ratings were each given a weight of "7," designating them as most pathological; the six having the next highest totals a weight of "6" and so on, with decreasing weights representing diminishing pathology. The items and their pathology scores are:

Pathology score of 7 (3 items)
- Flight of ideas
- Delusional thought
- Thought processes are regressive, archaic, primitive or magical

Pathology score of 6 (6 items)
- Inappropriate affect
- Metaphysical or philosophical ponderings of a vague unproductive nature
- Grandiose
- Withdraws into autistic fantasy
- Confused, illogical thinking
- Erotic gratifications regressive

Pathology score of 5 (8 items)
- Prominent mood swings
- Disregards social canons
- Intrusion of irrelevant material in thinking
- Projection an important defense
- Suspicious and wary
- Inadequate reality testing
- Free anxiety; intense and sudden in appearance with controls failing
- Maladaptive externalized aggression

Pathology score of 4 (10 items)
- Socially withdrawn
- Primarily identified with opposite sex
- Aggressive, destructive preoccupations
- Interest is dominated by one or two themes to the exclusion of all others
- Impulsive, unstable in reaction to usual life stimuli
- Intellectual functioning significantly below potential
- Denial an important defense
- Sexual preoccupation
- Undoing an important defense
- Depression

Pathology score of 3 (8 items)
- Blocking
- Generalized feelings of inadequacy and ineffectuality
- Isolation an important defense
- Guilt feelings
- Anxiety provoking conflicts expressed in conversion symptoms or psychosomatic illness
- Infantile and egocentric
- Anxiety persistent and pervasive
- Obsessive-compulsive behavior

Pathology score of 2 (6 items)
- Repression an important defense
- Displacement an important defense
- Unrealistic self-concept and level of aspiration
- Reaction formation an important defense
- Passive and submissive
- Characterized by rigidity

Pathology score of 1 (3 items)
- Rationalization an important defense
- Overconventional
- Limited insight

A convenient arrangement is to note the appropriate pathology score on the back of each clinical formulation card. As employed in personality evaluation, this weighted Q array is sorted in the ordinary manner. The clinician sorts each of the clinical formulations into one of seven categories, depending on how characteristic the description is of the patient. The number of items to be placed in each category is predesignated, so that an approximate normal distribution, identical with the one used for the pathology sort, is obtained. Items in Category 7 are those judged most applicable to the patient, while the formulations in Category 1 are considered entirely inappropriate descriptions.

Denoting the category in which a particular clinical formulation is placed in a personality assessment as its "applicability rating," a composite adjustment index can be computed for a patient by simply multiplying each applicability rating by the item's pathology score and summing the 44 products. If the clinical formulations indicative of the most severe pathology are always rated as most characteristic of the patient, then a maximum total will be obtained, reflecting the

greatest degree of disturbance. On the other hand, if the items judged as most pathological are consistently scored as least characteristic of the patient, the total is minimal, indicating relatively insignificant emotional disturbance. The highest possible total for the 44 clinical formulations using this procedure is 722, while the lowest total is 586. A range of 0 to 236 can thus be established for the adjustment index by subtracting the constant 586.

Reliability

As a check on the reliability of the adjustment index, weighted *Q*-sort personality evaluations done independently on the same patients by each author were contrasted. These sorts were based on the test protocols of 25 patients not tested by either author and selected at random from the files. In general, the battery of tests available on each patient consisted of the Wechsler, Rorschach, Thematic Apperception Test, Draw-a-Person Test, and the Sentence Completion Test. The correlation between the two series of adjustment indices, calculated from the two independent *Q* sorts for each patient, was .79, which is significant at .001 level, and indicates that this procedure can be used in consistent fashion.

Validity

To determine the validity of this procedure, adjustment indices from the two *Q*-sort ratings for each of the 25 patients were averaged, and the mean ratings were compared with estimates of adjustment obtained from each patient's therapist. Each of the psychiatrists involved was asked to rate the patient as he was at the time of the psychological testing, using The Institute of Living Clinical Rating Scale for Attending Physician (**5**). This scale consists of 12 items, each to be rated on a five-point scale of increasing maladjustment. The ratings may be summed to give an over-all estimate of adjustment, with a maximum total score of 60 corresponding to maximum maladjustment. The total scores from the therapist rating were found to be correlated with the mean *Q*-sort indices .57, a result which is significant at the .01 level. Correcting for attenuation increases this correlation coefficient to .66.

Discussion

This modified procedure has a number of advantages compared with the conventional correlation approach. First, the computation of the adjustment index is somewhat less arduous than the derivation of a correlation, although the latter computation can be simplified by use of the "method of differences" (2, 4). Also, the adjustment index escapes certain disadvantages inherent in the *r*: the potentiality for being mistaken for a percentage, and the complexity of interpretation of a measure that may vary in significance.

But the most important advantage of the adjustment index approach lies in its ability to compare *Q* sorts from *groups* of patients, and in its amenability to further relatively simple statistical manipulation. Returning to the example of contrasting "unweighted" *Q*-sorts personality evaluations done on a number of patients prior to and after therapy, the method of comparison is a correlation, calculated separately for each patient. The researcher is thus left with a distribution of *r*'s which may be exceedingly difficult to interpret. When the weighted *Q* sort is used, however, pre- and posttherapy adjustment indices are computed for every patient in the manner described previously. A statistical test is performed next to determine if the mean difference between the two series of adjustment indices is significant. This procedure then gives a definitive picture of the scope and direction of changes associated with therapy.

The advantage of the index is even clearer in the case where two groups of patients are to be compared before and after therapy. The investigator using correlations would be faced with the task of comparing two series of *r*'s, the before-after therapy correlations for the two groups. The *r*'s cannot be compared directly, but must be converted to *z* values; whereas two series of adjustment scores differences could be readily compared by conventional tests of significance. Furthermore, the sign of the index difference immediately indicates the direction of change, while direction cannot be discerned from the *r* comparison.

The weighted *Q* sort also readily lends it-

self to a comparative analysis of the ratings of psychiatrist and psychologist judges, in that the 44 clinical formulations can be rated from the results of psychological tests or therapeutic interviews. Of particular interest would be the use of weighted *Q* sorts to determine the degree of agreement between psychologists and psychiatrists evaluating the general adjustment of the same group of patients.

The authors' experience with the weighted *Q* sort in its present form suggested that augmenting the 44 clinical formulations with additional items would enhance the sensitivity of the technique.

Summary

A method was described for modifying the *Q* sort so that it could be utilized to estimate over-all degree of emotional disturbance and also to determine extent of personality change. Briefly, the procedure involved weighting each of 44 *Q*-array clinical formulations according to how indicative the specific item was of serious pathology. The statements can then be sorted, in the usual manner, into categories arranged on continuum: personality descriptions judged most characteristic of the patient to be placed at one end of the continuum, while those rated as least characteristic at the other extreme. From the relative applicability of the various clinical formulations to the particular patient and the assigned weights of the items, an individual "adjustment index" is determined. Reliability and validity data with reference to this procedure are presented and the advantages of using the weighted *Q* sort's adjustment indices in group comparisons are discussed.

Received January 9, 1958.
Early Publication.

References

1. Beck, S. J. *The six schizophrenias.* New York: American Orthopsychiatric Association, 1954.
2. Cohen, J. An aid in the computation of correlations based on Q sorts. *Psychol. Bull.*, 1957, **54**, 138–139.
3. Dymond, Rosalind F. Adjustment changes over therapy from self-sorts. In C. R. Rogers & Rosalind F. Dymond (Eds.), *Psychotherapy and personality change.* Chicago: Univer. of Chicago Press, 1954. Pp. 76–84.
4. Hilden, A. H. *Manual for Q-sort and random sets of personal concepts.* St. Louis, Mo.: Author, 1954.
5. Reznikoff, M., & Zeller, W. W. A procedure for evaluating the status of schizophrenic patients. *J. clin. exp. Psychopath. & Quart. Rev. Psychiat. Neurol.*, 1957, **18**, 367–371.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Adjustment and Prediction of Others' Behavior

June Elizabeth Chance
University of North Carolina

Among the myriad problems in the area of interpersonal perception are those concerning the characteristics of the individual perceiver which affect the adequacy of his perceptions. Implicit in many conceptions of maladjustive behavior is the notion that the poorly adjusted individual is less able to perceive clearly and hence to predict the responses of others. However, empirical evidence regarding the relationship between personal adjustment and ability to predict is equivocal (5). A recent study by Bieri, Blacharsky, and Reid (1) hypothesized a negative relationship between degree of maladjustment and accuracy of prediction, as well as a positive relationship between degree of maladjustment and tendency to engage in projection (assuming similarity between self and others). Neither hypothesis was supported by the findings; in fact, the obtained relationship between degree of maladjustment and projection was in the direction opposite from that predicted. On the basis of their study, these investigators concluded that "better adjusted *Ss* tend to predict most accurately on the basis of *similarities* between themselves and others, while maladjusted *Ss* emphasize to a greater degree than adjusted *Ss* accurate prediction of differences between themselves and others" (1, p. 353).

Some methodological aspects of this study suggest, however, an alternative explanation of the findings as has been recently pointed out by Erikson (3). The choice of the Taylor Manifest Anxiety Scale as the set of responses to be predicted leaves open the possibility that *Ss* in making their predictions could have been responding to a stereotype of how a college student responds. More specifically, if most *Ss* were aware for most items of the MAS that certain responses were more "acceptable" and probably more likely to occur in a college population, the results obtained by Bieri, Blacharsky, and Reid would automatically follow. An *S* who himself had given maladjusted responses would "predict" the responses of his target figures to be mostly different from his own and consequently a high proportion of his accurate predictions would be perceived differences. Conversely, the well-adjusted *S* responding in the more acceptable and likely manner for himself and also responding in this manner for his targets would obtain most of his accuracy in perceived similarities. Further, in this study *Ss* chose members of their class to predict, and since differential acquaintance with fellow class members might be related to the variable of adjustment, this procedure could easily confound the results of the study.

The present study is an attempt to test the hypothesis that there is a negative relationship between personal adjustment and adequacy of performance in predicting the behavior of others. The method used somewhat obviates the possibility that *Ss* can predict on the basis of a stereotype and also the difficulties which may ensue in interpreting results where *Ss* select acquaintances to predict.

Method

Fifty-two *Ss* from two sections of an elementary psychology course at the University of North Carolina made predictions for their respective section instructors. Both groups were administered the Rotter Incomplete Sentences Blank (4) and a modified group form of the Minnesota Multiphasic Person-

ality Inventory, which included the Welsh *A* [1] and *R* Scales (7). At a later time, after approximately 35 hours of class meetings, the prediction procedure was administered. *Ss* were asked first to respond to the prediction schedule for themselves and immediately after to predict how their respective instructor had responded to the same items. All experimental procedures were administered by an experimenter [2] who had nothing to do with the teaching of the course, and *Ss* were assured that their instructors would not see the data until after the end of the term.

The criterion instrument on which the predictions were made was constructed by selecting 40 items from the Edwards Personal Preference Inventory (2). The inventory is an attempt to measure psychological needs by means of a forced choice procedure. Each item consists of a pair of statements from which the individual is to select the one most descriptive of him. Each pair of statements is equated for social acceptability; that is, each member of a pair is about equally likely to be chosen in a college population. The 40 item pairs used for the prediction schedule included 10 pairs each from the scales measuring needs for achievement, affiliation, aggression, and autonomy. The choice of items was a priori, but with two objectives—to provide a variety of behaviors to be predicted and behaviors of a sort that *Ss* would be likely to possess some relevant observations from which to make their inferences.

The target figures were the instructors of the two sections: one instructor was male, one female; each section predicted its own instructor. The sections met at the same hour and one half of the students enrolled in the course had been arbitrarily assigned to each. Both instructors had taken the Edwards' PPI out of curiosity about six months before the present study was suggested.

Scores for predictive behavior were derived in the same manner as in the Bieri, Blacharsky, and Reid study. Predictive accuracy scores were obtained by totalling the number of agreements between *S*'s predictions and the target's responses. Assimilative projection scores were obtained by totalling the number of *S*'s predictions (accurate and inaccurate) which coincided with his own responses. The latter score was separated into two components: accurate projection, i.e., accurately perceived similarities; and inaccurate projection. A perceived difference score was obtained by totalling the number of disagreements between *S*'s predictions and his own responses. This score was also separated into two parts: accurately and inaccurately perceived differences. In addition, the number of real similarities and the number of real differences between each *S* and the target were totaled.

Results

The Welsh scales and ISB yielded distributions of scores very similar to those reported by the authors for other college populations. Two judges independently scored 28 of the 52 ISB protocols; the obtained interscorer reliability was .89. The correlation between the Welsh *A* scale and the ISB was .38; between the *R* scale and the ISB, .30; between *A* plus *R* and the ISB, .35. All of these correlations among the various measures of adjustment are significant beyond the .05 level. (The correlation of .04 between the *A* and *R* scales for this group is not significantly different from zero.)

The prediction task given the *Ss* was a difficult one. The mean predictive accuracy score for all *Ss* was 22.39 (*SD* 1.99). Since a predictive accuracy score of 20 could presumably be obtained by chance, a test of the significance of the difference between the obtained mean and 20 was performed (6, p. 147). The value of t equalled 7.68 ($p < .001$, df 51).

The product-moment correlations between the various prediction scores and the measures of adjustment are presented in Table 1. Fisher's z transformation was used to test the

[1] Welsh suggests that the *A* scale measures an aspect of maladjustment characterized by anxiety and dysphoria. Mr. Carl Cochran, in a doctoral dissertation in progress, found a correlation between the MAS and the Welsh *A* scale of + .81 for a group of 514 college students. The *R* scale was included to get at another aspect of adjustment—tendency toward denial and repression.

[2] The author wishes to thank Howard Leventhal for his assistance in the collection of these data.

Table 1

Relationships Between Measures of Adjustment and Prediction Scores ($N = 52$)

Measure	ISB	*A* scale	*R* scale
Predictive accuracy	−.26*	.02	−.14
Assimilative projection	.05	−.28**	.02
Accurate projection	.00	−.18	.05
Inaccurate projection	.11	−.26*	.21
Perceived differences	−.05	.28**	−.02
Accurately perceived differences	−.37***	.16	−.02
Inaccurately perceived differences	.16	.19	−.02

* Indicates $p < .10$.
** Indicates $p < .05$.
*** Indicates $p < .01$.

significance of the differences between the obtained *r*'s and zero. For all the measures of adjustment a high score is indicative of maladjustment.

The correlation between predictive accuracy and ISB scores although in the expected direction did not quite reach statistical significance ($.10 > p > .05$). Neither *A* nor *R* correlated significantly with predictive accuracy.

While ISB scores do not correlate with tendencies to assume similarity, either accurately or inaccurately, scores on the *A* scale do seem to be negatively related to inaccurate projection (assumption of unwarranted similarity), meaning that the higher the *A* score the less likely the person to assume unwarranted similarity between himself and the person predicted for. Although the positive correlation of *R* with inaccurate projection does not quite reach the .10 level, it suggests that the independent aspects of adjustment measured by the *A* and *R* scales are related to tendencies "to see similarities where they are not" in quite different ways. Also inspection of the relationships between the three projection scores and the *A* and *R* scales suggests that while the negative relationship between *A* and inaccurate projection could well be largely the result of some general response set to assume dissimilarity, the positive relationship between *R* scores and inaccurate projection might be independent of such a set.

ISB scores are not significantly related to over-all perception of differences between self and the target predicted; however, they are negatively related ($p < .01$) to accurate perception of differences, suggesting that maladjusted *S*s predict differences less well than do better adjusted *S*s. In contrast, scores on the *A* scale are positively related to over-all perception of differences, although neither the accurate nor inaccurate components of these predicted differences correlate significantly with scores on the *A* scale.

Real differences between *S* and the predicted target correlated with predictive accuracy − .07, which is not significant. Real differences correlated + .02 with scores on the *A* scale and − .16 with scores on the ISB. Accurately predicted differences correlated with real differences + .64. When a partial correlation between accurately predicted differences and ISB scores holding real differences constant is computed the value is − .36.

Discussion

The above findings do not support Bieri, Blacharsky, and Reid's conclusions, i.e., that better adjusted *S*s are more likely to predict similarities accurately, while more poorly adjusted *S*s are more likely to predict differences accurately. The obtained results do suggest that there may be a general response set to assume similarity or difference when making predictions related to the dimension of adjustment measured by the *A* scale, but not related to the dimensions measured by either the ISB or the *R* scale. Maladjustment as measured by the *A* scale seems to be related to a tendency to believe that others are different from one's self; hence, in a study where the targets predicted are truly quite different from the *S*s, spurious positive correlations between maladjustment and predictive accuracy could be obtained.

The hypothesis of a negative relationship between adjustment and adequacy of predictive behavior is neither clearly refuted nor confirmed. The findings demonstrate a trend toward a negative relationship between ISB scores and over-all accuracy of prediction in the kind of experimental situation described here; this negative relationship is most apparent in the differential ability among indi-

viduals of different levels of adjustment to predict differences between self and others accurately.

Generalizations of these findings are limited by the fact that the three measures of adjustment used, while positively correlated among themselves, are related to the predictive accuracy scores in different ways. Statements about the relationship of adjustment to prediction behavior are not meaningful apart from the particular criteria of adjustment employed. Although the study as executed has the virtue that all *Ss* had rather extended and fairly uniform acquaintance with the targets predicted, it must be pointed out that most other studies in this area have selected target figures who were either *S*'s peers or who could readily be identified by *S* as peers. To the degree that *Ss* might be expected to operate with a different set of assumptions about figures approximately 10 years older than they and with Ph.D.'s in psychology, the results of the present study may not be comparable to those of other studies.

Perhaps better answers in this exceedingly complex area could be obtained if the questions asked were framed differently. That is to say, if questions were asked not in terms of highly abstracted variables like adjustment-maladjustment, but rather in terms of what aspects of adjustment or what particular and characteristic modes of dealing with problems might be related to adequacy of interpersonal perceptions. Further studies could also profit from more specification of what is predicted. One might hypothesize that given modes of adjustment are differentially related to ability to predict certain classes of behavior on the part of others. For example, individuals who have difficulty in coping with their own aggressive propensities might be expected to be less accurate in predicting aggressive responses on the part of others, but relatively more accurate in predicting others' dependent responses.

Summary

Fifty-two college students were asked to respond to a 40-item, forced choice questionnaire and then to predict their respective instructor's responses to the same questionnaire. On a previous occasion, three paper-pencil measures of adjustment had been obtained for these *Ss*—the Rotter Incomplete Sentences Blank, and the Welsh *A* and *R* scales (MMPI). Scores for over-all accuracy of prediction, various components of accuracy, and tendencies to assume similarities or differences between self and others in making predictions were correlated with the measures of adjustment. Results obtained indicate a trend for adjustment as measured by the ISB to be negatively related to adequacy of prediction. Scores on the *A* scale were negatively related to tendency to assume similarity in making predictions.

Received June 3, 1957.

References

1. Bieri, J., Blacharsky, E., & Reid, J. W. Predictive behavior and personal adjustment. *J. consult. Psychol.*, 1955, **19**, 351–356.
2. Edwards, A. L. *Manual: Edwards Personal Preference Schedule.* New York: Psychological Corp., 1954.
3. Erikson, C. W. Personality. *Annu. Rev. Psychol.*, 1957, **8**, 185–210.
4. Rotter, J. B., & Rafferty, Janet E. *Manual for the Rotter Incomplete Sentences Blank, College Form.* New York: Psychological Corp., 1950.
5. Taft, R. The ability to judge people. *Psychol. Bull.*, 1955, **52**, 1–23.
6. Walker, H. M., & Lev, J. *Statistical inference.* New York: Holt, 1953.
7. Welsh, G. S. Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: Univer. of Minnesota Press, 1956.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Nonprojective Determinants of Perception on the TAT[1]

Bernard I. Murstein
Louisiana State University

Projective techniques are coming under ever closer scrutiny with regard to the meaning of the material elicited by these techniques. Since the various techniques differ in their stimulus properties (clay, ink blots, dolls, scenes of people, etc.), it would seem of importance to investigate the nonprojective determinants of perception as well as the more personal aspects, in order to interpret the content in a more meaningful manner.

Eron in earlier studies (1, 2) showed that the emotional tone of the stories told to the TAT cards was more often negative (unhappy, sad, hostility, death) than positive (pleasant, friendly, contented).

It should follow, therefore, that a positive correlation ought to exist between the emotional tone and the positive-negative stimulus value of the cards such that negative scenes tend to elicit negatively toned stories.

In addition, clinical experience with the TAT cards has suggested to the author that in those cards which are relatively unstructured or ambiguous with regard to their stimulus properties, positively-toned stories were often elicited. Conversely, clearly structured or unambiguous cards often elicited negatively-toned stories. The last observation does not necessarily imply any innate relationship between structure and tone, but rather suggests that in their selection of cards in constructing their tests, Murray and Morgan chose many highly structured cards whose stimulus properties also were quite negative with regard to the emotions expressed.

The following hypotheses were therefore formulated:

1. The emotional tone of stories told in response to the TAT cards is positively related to the pleasantness-unpleasantness of the TAT cards. Those cards pleasant in appearance yield pleasantly toned stories.

2. The emotional tone of stories is also a function of the clarity of the emotions expressed on the TAT cards. Those cards most clearly structured (unambiguous) will yield negatively toned stories while unstructured cards yield more positive themes.

3. While some overlap exists between the dimensions of pleasantness-unpleasantness, and ambiguous-unambiguous, both dimensions show sufficiently low intercorrelation to be considered separately as factors determining the tone of the stories told.

Procedure

Twelve normal female subjects (*Ss*) (8 students, 4 secretaries), ranging in age from 18 to 39, with a median age of 23, were given, in a randomized fashion, the same 20 TAT cards that Eron (2) used in his earlier study with women (1, 2, 3 GF, 4, 5, 6 GF, 7 GF, 8 GF, 9 GF, 10, 11, 12 F, 13 MF, 14, 15, 16, 17 GF, 18 GF, 19, 20). The age range and median were roughly equivalent to that of Eron's group.

The emotional tone ratings quantified by Eron on a five-point scale ranging from 2 to −2 according to their pleasantness-unpleasantness were totaled up for each card from his study, omitting rejects. These sums were then converted into ranks so that the most unpleasantly toned card was ranked 1, while the most pleasant card was ranked 20.

The 20 cards were given individually by four different experimenters to the 12 *Ss*,

[1] Paper read at the meeting of the American Psychological Association, New York, 1957.

with the author administering the cards to six *Ss*, while the others administered the cards to two *Ss* each.[2] Essentially, the following instructions were given:

> I want you to rank each card according to how clear the card is with regard to the emotions expressed in it. It does not matter whether the card is happy, sad, bored, friendly, etc. That card in which the emotion expressed is most easily determined you will rank as 1. That card which is next clearest in terms of some definite emotion expressed, rank as 2, and so forth, until the card ranked as 20 should be the one which because it is so ambiguous is capable of being seen as expressing many possible emotions.

The instructions for ranking according to the pleasant-unpleasant dimension were:

> I want you to rank each card according to how unpleasant or unhappy the emotion expressed seems to be. That card which is the most unpleasant or unhappy should be ranked 1, the card next most unpleasant should be ranked 2, etc., until 20 should represent the card which is the *least* unhappy or unpleasant, or more simply, that card which is the most pleasant one.

Half of the *Ss* received the clarity-lack-of-clarity task first, while the other half received the pleasant-unpleasant ranking first. Subsequent appointments were made by the examiners for the "second part" of the experiment, which occurred from three days to three months later. The "second part" was a retesting for reliability purposes.

Results

The rho coefficients between the Eron emotional tone-of-story dimension from his study, and in this study the ambiguous-unambiguous, and pleasantness-unpleasantness dimensions of the stimulus properties of the cards are indicated for each *S* in Table 1. This table also shows the average rho intercorrelation between these dimensions for the pooled subjects. Test-retest reliability for the 10 of the 12 *Ss* available for retesting, covering an interval of from three days to three months, was .91 for the pleasantness dimension and .84 for the ambiguity dimension. The reliability was not significantly related to the duration of time between the first and second administration of the cards. Individual variation between the 12 *Ss* in the correlation of the ambiguous-unambiguous dimension with emotional-tone ranking is considerable (χ^2; $p < .10$) though not for the pleasant-unpleasant vs. emotional-tone correlation (χ^2; $p < .80$). Both of the dimensions (pleasant-unpleasant, and ambiguous-unambiguous) are significantly related to the pooled *Ss*' emotional-tone rankings as taken from Eron's study beyond the .01 point for a one-tailed test of significance. The respective correlations are .59, and .56. Hypotheses 1 and 2, therefore, are confirmed. The average correlation between the pleasantness-unpleasantness and ambiguous-unambiguous dimensions is significant at the .01 point with rho = .40. By use of the formula for correction for attenuation, it may be noted that 45% of the "true" variance of the emotional tone ratings is accounted for by the pleasantness-unpleasantness vs. emotional tone-rating correlation, and 44% of the "true" variance is accounted for by the ambiguous-unambiguous vs. emotional-tone rating correlation. The "true" common variance between pleasantness-unpleasantness and ambiguity-unambiguity is only 21%. Hence, both of these dimensions

Table 1

Correlations Between Pleasantness, Ambiguity, and Emotional Tone Dimensions for Each *S* and for Pooled *Ss*

S	Ambiguity vs. Emotional Tone	Pleasantness vs. Emotional Tone	Ambiguity vs. Pleasantness
1	.58	.60	.69
2	.53	.29	.02
3	.22	.56	−.12
4	.58	.50	.60
5	.75	.60	.59
6	.16	.60	−.12
7	.53	.65	.17
8	.52	.50	.21
9	.83	.73	.75
10	.74	.75	.68
11	.23	.70	.09
12	.60	.44	.64
Avg. r	.56*	.59*	.40*

* Significant beyond .01 level.

[2] The author expresses his thanks to Deldon McNeely, Ronald Pryer, John Roitzsch, and Richard Rolston for their aid in administering the TAT cards.

must be taken into account in predicting the emotional tone ratings to TAT cards. Therefore, Hypothesis 3 is seen to be confirmed.

Lastly, neither order of presentation of the dimensions nor age was found to be related to the size of the correlation for any dimension.

Discussion

It would be helpful in evaluating the findings of this study to utilize a schema of Helson's theory of adaptation-level (3). The theory briefly stated would hold that perception is a resultant of the pooled interaction of the stimulus, background, and residual factors (i.e., personality). Singer (7) has stated that clinicians have been too prone to overweight the importance of personality in determining perception, to the neglect of stimulus and background factors. The results of this experiment seem to support Singer in confirming the existence of several nonprojective determinants of perception on the TAT. The stimulus-determined dimensions of pleasantness and ambiguity together accounted for 61% of the true common variance with the emotional tone-of-stories dimension. Presumably, within the remaining true variance the mechanism of projection occurs, though this mechanism itself has been found to be exceedingly complex (4, 5, 6).

The findings of this study, therefore, would seem to call for a bridling of the unrestrained assumption that, for example, perception of hostility on one of the TAT cards is necessarily indicative of a "hostile" personality. In a technique such as the TAT where the stimulus properties vary considerably in ambiguity and pleasantness from card to card, the perception of hostility on one card might indicate considerable distortion of the stimulus properties (projection?), while on another card reflect only considerable respect for the objective stimulus.

A clinician reviewing the themes of the stories without noting the stimulus characteristics of the card that occasioned it, might draw, therefore, entirely inaccurate inferences. Consequently, it might be beneficial to construct a scale of cards whose pictures differed in pleasantness, structure, etc., by some measured amount based on the construction of judged group norms. In this way, clinicians might come closer to a realization of the suspected correlation between "deviation of perception" and deviation of personality. In short, "deviation of perception" must take cognizance of measured stimulus and background characteristics of the TAT cards before we attribute the remaining variance to deviations in personality patterns.

Summary

The purpose of this study was to examine some of the nonprojective determinants of perception on the TAT.

Three hypotheses were formulated:

1. The emotional tone of stories told in response to the TAT cards is positively related to the pleasantness-unpleasantness of these cards. Those cards pleasant in appearance yield pleasantly toned stories.

2. The emotional tone of the stories is also a function of the structure of the emotional expression on the cards. Those cards most clearly structured (unambiguous) will yield negatively toned stories, while unstructured (ambiguous) cards yield more pleasant themes.

3. While some overlap exists between the dimensions of pleasantness-unpleasantness and ambiguity-nonambiguity, both dimensions show sufficiently low intercorrelation to warrant consideration as separate factors determining the tone of the stories told.

A rho coefficient of .56 between ambiguity-nonambiguity and emotional tone, and .59 between pleasantness-unpleasantness and emotional tone (both significant at the .01 point) seemed to lend support to Hypotheses 1 and 2.

Hypothesis 3 found confirmation in that while both of the dimensions of pleasantness-unpleasantness and ambiguity-nonambiguity correlated $r' = .40$ with each other, each had considerable valid variance in common with the emotional tone dimension which was not accounted for by this correlation.

Received June 14, 1957.

References

1. Eron, L. D. A normative study of the Thematic Apperception Test. *Psychol. Monogr.*, 1950, **64**, No. 9 (Whole No. 315).

2. Eron, L. D. Responses of women to the Thematic Apperception Test. *J. consult. Psychol.*, 1953, **17**, 269–282.
3. Helson, H. Adaptation level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, **55**, 297–313.
4. Murstein, B. I. The projection of hostility on the Rorschach, and as a result of ego-threat. *J. proj. Tech.*, 1956, **2**, 418–428.
5. Murstein, B. I. Some comments on the measurement of projection and empathy. *J. consult. Psychol.*, 1957, **21**, 81–82.
6. Murstein, B. I. Studies in projection: A critique. *J. proj. Tech.*, 1957, **21**, 129–136.
7. Singer, J. L. The non-projective aspects of the Rorschach experiment: V. Discussion of the clinical implications of the non-projective aspects of the Rorschach. *J. soc. Psychol.*, 1956, **44**, 207–214.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

The Validity of the Objective Rorschach and the MMPI[1]

William D. Winter
San Jose State College

and Ramon A. Salcines
Milwaukee County Guidance Clinic

Because of heavy demands for their diagnostic services, psychologists have attempted to develop valid personality tests which are quickly scored, easily evaluated, and adaptable to testing large groups of subjects (*Ss*). A number of such rapid screening devices are now in use. One of the newest to be offered to the busy clinician is the Objective Rorschach, a variation of the multiple choice Rorschach, developed by P. O. O'Reilly (1, 3, 5). In this procedure, the *S* is tested individually and he is shown each Rorschach card in turn. Together with each card, he is presented with a list of 12 responses and is asked to choose the two responses which he feels best describe the card. His choices are scored and an over-all measure of disturbance is derived by adding the scores obtained on each card. O'Reilly's data indicate that this procedure is an effective device for distinguishing among normals, neurotics, and psychotics (5), and that the over-all score is also significantly related to clinical improvement in response to therapy (6). The purpose of the present study is to test the validity of the Objective Rorschach when used with a different sample of *Ss*, and to compare its ability to identify mental patients with that of a more conventional test, the MMPI (2).

[1] These data were gathered by the second author as part of a master's thesis undertaken at the University of Detroit. The authors wish to express their gratitude to the staffs of the various hospitals and schools, and to the social organizations who cooperated in this study by generously contributing their time and advice.

Method

The *Ss* for this study consisted of three groups of white males, matched closely as to age and education. Their age range was 17–45 (median 26.6), and their education ranged from 7–17 years of school completed (median 12.2). The first group consisted of 25 normals who were functioning adequately in the community, and who denied ever having had any serious personality difficulties. The second group consisted of 25 normal *Ss* who were hospitalized with an acute physical illness, an injury or infection, which as far as could be determined was not obviously psychosomatic in origin. These *Ss* were in a stressful situation and experiencing physical and mental discomfort. The third group consisted of 25 hospitalized schizophrenics, all of whom were testable but, in the judgment of the staff, grossly disturbed in their thinking or behavior.

Each *S* was tested by the same psychologist. The test battery included the Objective Rorschach and the group form of the MMPI. The measures of disturbance used in this study are the total Objective Rorschach score and the total score obtained on the MMPI psychosis scale developed by Peterson (7). In scoring this latter scale, the *S* is given one point for each of the following psychotic signs appearing in his MMPI, and the number of points are cumulated:

1. *T* scores over 70 on four or more of the clinical scales
2. *F* greater than 65 (*T* score)
3. *Sc* greater than *Pt*
4. *Pa* or *Ma* greater than 70 (*T* score)
5. *Pa* or *Sc* or *Ma* greater than *Hs* and *D* and *Hy*
6. *D* greater than *Hs* and *Hy*

Peterson suggests that the presence of three or more of these signs indicates psychosis. The ability of the MMPI psychosis scale and the Objective Rorschach to differentiate among our three groups of *Ss* was tested by means of chi square.

In analyzing the results of this study, the major emphasis was placed on the ability of these two personality tests to predict psychosis. Specifically, we were interested in whether or not the published cutoff scores on these tests succeeded in separating our schizophrenic patients from the two nonpsychotic groups. In evaluating the predictive efficiency of these tests, it should be kept in mind that only one-third of our sample of 75 *Ss* consists of psychotics. Therefore, if one were to arbitrarily diagnose each *S* as not being psychotic, two-thirds of the diagnoses would be correct. For a test to be of value in this situation, its use would have to allow accurate diagnosis of more than the 50 *Ss* who could be correctly diagnosed without knowledge of their test results. In addition to the statistical questions of whether or not the three groups obtained significantly different (median) scores on these two tests and whether or not using these tests leads to more correct diagnoses than can be obtained without the tests, one must take into consideration the practical problem of whether the additional number of correct diagnoses one might obtain justifies the time and trouble it takes to administer a given test.

Results

Objective Rorschach

Table 1 shows the distribution of Objective Rorschach scores obtained by our three groups. In his original article, O'Reilly (5) recommends using a cutoff score of 38 and above as an index of psychosis. Although the use of this cutting point leads to the correct classification of 73% of our *Ss*, examination of Table 1 will show that only five of our 75 *Ss* obtained scores above 37. Although these five were all in the schizophrenic group, 80% of the schizophrenics fall below this cutting point, which would thus appear to be set too high to be of practical value for our sample. A median test showed that our schizophrenic *Ss* scored significantly lower on this test than did O'Reilly's psychotic *Ss* ($\chi^2 = 18.0$, 1 *df*, $p < .001$). Since O'Reilly does not describe the composition of his sample, we can offer no hypotheses as to why this large difference occurred. Although a higher percentage of correct diagnoses was made by using the test than could be made without it, this increase is insignificant, and the meager results obtained scarcely justify the effort involved in individually testing 75 *Ss*.

Further examination of Table 1 will show that a cutoff score of 36 and above, which also correctly classifies 73% of our *Ss*, is also of little practical value, since the addition of five correct diagnoses over the 50 obtainable without using the test is not a significant increase. The use of this cutoff results in the correct diagnosis of only nine psychotics, at the expense of incorrectly labelling four non-

Table 1

Distribution of Objective Rorschach Scores

Score	Normal	Physically ill	Psychotic
49			1
43			1
42			1
41			0
40			0
39			2
38			0
37	0	3	0
36	0	1	4
35	1	3	0
34	3	3	1
33	2	1	5
32	2	0	3
31	1	2	2
30	4	2	3
29	1	4	0
28	5	3	0
27	3	1	0
26	3	1	1
25	0	1	1
N	25	25	25

Table 2

Ss Falling Above and Below the Objective Rorschach Common Median

Score	Normal	Physically ill	Psychotic
32–49	8	11	18
25–31	17	14	7

$\chi^2 = 8.4$; $df = 2$; $p < .02$.

Table 3

Ss Obtaining Scores of Three or Above on MMPI Psychosis Scale

Score	Normal	Physically ill	Psychotic
3–6	3	5	19
0–2	22	20	6

$\chi^2 = 26.4$; $df = 2$; $p < .001$.

psychotics as being psychotic. Since the two cutting scores discussed above result in the highest percentage of correctly classified *Ss*, and since neither proves to be satisfactory, we must conclude that the O'Reilly Objective Rorschach is valueless for the practical problem of making individual diagnoses of psychosis in our *Ss*.

To determine whether the Objective Rorschach differentiated our three groups, an extension of the median test was performed (8). Table 2 shows the number of *Ss* in each group falling above and below the common median of the entire sample of 75. In this case, the cutoff score is 32 and above, and chi square for this contingency table is 8.4, which for two degrees of freedom is significant beyond the .02 level. Therefore, we can reject the hypothesis that our three groups come from the same population or from populations with equal medians. The Objective Rorschach does have some measure of validity in that the median scores of the three groups differ from one another, although the relationship between test scores and diagnostic category is not high enough to justify the use of this test in making individual predictions of psychosis in a sample such as ours, in which one-third of the *Ss* are psychotic. Whether the forecasting efficiency of this test can be greatly improved by item analysis must await further study.

MMPI Psychosis Scale

Peterson's results (7) suggest that a score of three and above on his scale is a conservative yet effective indicator of psychosis. Table 3 shows the number of *Ss* in each group who obtained scores above and below this cutting point. By using this cutoff, 81% of our *Ss* are correctly classified as to whether or not they are psychotic. This represents a significant increase (22%) of 11 correct diagnoses over the 50 that could be correctly determined without the test.

As with the Objective Rorschach, a median test was performed to see whether the medians of the three groups differ. Table 4 presents the number of *Ss* in each group who scored at or above the common median (two and above) of the entire sample. The chi square for this table is 18.2 ($p < .001$). Therefore, Peterson's psychosis scale on the group-administered MMPI is more effective than the individually administered Objective Rorschach in differentiating our three groups and in making individual diagnoses of psychosis. However, since a one-point difference in the cutoff score on the Peterson scale makes a distinct difference in the number of correct diagnoses, one might well question the validity of judgments made on the basis of such a short instrument, especially when a given clinical patient obtains a score at or near the cutting point. It should also be noted that both the Peterson scale and the Objective Rorschach drew the majority of their "false positives" from the group of nonpsychotic, physically ill *Ss*.

Table 4

Ss Falling Above and Below the MMPI Psychosis Scale Common Median

Score	Normal	Physically ill	Psychotic
2–6	6	14	21
0–1	19	11	4

$\chi^2 = 18.2$; $df = 2$; $p < .001$.

Table 5

Correlations Between Objective Rorschach Score and MMPI Scales

MMPI scale	*r*
K	−.09
F	.37**
Hs + .5*K*	.22
D	.40**
Hy	.25*
Pd + .4*K*	.29*
Mf	−.04
Pa	.30**
Pt + 1*K*	.37**
Sc + 1*K*	.41**
Ma + .2*K*	.33**
Psychosis scale	.44**

* $p < .05$.
** $p < .01$.

In Table 5 are listed the product-moment correlations between the Objective Rorschach total score and each of the MMPI scales for our entire sample of 75. There are several low, but significant correlations present, especially between the Objective Rorschach and those MMPI scales indicative of psychosis. The two tests are to some extent measuring the same things, but for the purposes investigated by this study, the MMPI psychosis scale seems to be a more effective instrument than the Objective Rorschach.

Summary and Conclusions

A sample of 25 normal, 25 physically ill, and 25 schizophrenic white male subjects was tested with the O'Reilly Objective Rorschach and the group form of the MMPI. The basic problem investigated is whether the Objective Rorschach or the MMPI psychosis scale developed by Peterson is more effective in differentiating the psychotic *Ss* from the other two groups. Our results indicate that although the three groups obtained significantly different medians on the Objective Rorschach, and although this test showed significant correlations with several of the MMPI scales, nevertheless the Peterson MMPI psychosis scale was much more effective in making individual predictions about whether or not a given *S* was psychotic. The Objective Rorschach was of no practical value for this purpose. For our sample, the MMPI psychosis scale was a more valid measure of psychotic disturbance. The MMPI cutoff score suggested by Peterson's results was applicable to our *Ss*, but the Objective Rorschach cutoff score suggested by O'Reilly was too high to be useful, since our psychotic *Ss* as a group scored significantly lower on this test than did O'Reilly's psychotic sample. Both tests drew most of their "false positives" from the group of physically ill patients. This suggested that a clinician should evaluate the stresses found in a person's life situation before attempting to make a diagnosis of psychosis on the basis of a score on either test.

Although the Objective Rorschach is an interesting and potentially useful screening device, in its present stage of development it does not appear to offer any advantage over the more commonly used and more easily administered and scored MMPI, at least for the purpose investigated in this present study. Further research is needed to cross-validate our findings and to ascertain why our psychotics differed from O'Reilly's sample. It would also be important to determine if the Objective Rorschach is of greater predictive value than tests like the MMPI in such areas as therapeutic improvement (6).

Received June 24, 1957.

References

1. Harrower-Erickson, Molly R., & Steiner, M. E. *Large scale Rorschach techniques.* Springfield, Ill.: Charles C Thomas, 1945.
2. Hathaway, S., & McKinley, J. C. *Minnesota Multiphasic Personality Inventory: Manual.* New York: Psychological Corp., 1951.
3. Kellman, S. Multiple choice Rorschach. In A. Weider (Ed.), *Contributions towards medical psychology.* New York: Ronald Press, 1953. Pp. 625–636.
4. Meehl, P. Profile analysis of the MMPI in differential diagnosis. *J. appl. Psychol.*, 1946, **30**, 517–524.
5. O'Reilly, P. O. The objective Rorschach: A suggested modification of Rorschach technique. *J. clin. Psychol.*, 1956, **12**, 27–31.
6. O'Reilly, P. O. Therapy: A proposed objective evaluation. *Dis. nerv. Syst.*, 1955, **16**, 1–4.
7. Peterson, D. R. The diagnosis of subclinical schizophrenia. *J. consult. Psychol.*, 1954, **18**, 198–200.
8. Siegel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Age and Sex Influences on MMPI Profile Peak Distributions in an Abnormal Population

Bernard S. Aaronson
New Castle State Hospital, New Castle, Indiana

Sex and age are factors which may have differential influence in bringing people into membership in a population of abnormals. In the case of sex, differences in morphology and physiology, as well as differences in life experiences, may all contribute to differing personality characteristics which cause males and females to be classified as abnormal. In the case of age, different personality factors may help create difficulty for individuals at various stages of the life process. To some extent, such differences may be a function of the demands made upon the individual by our culture as he moves through life. To some extent, particularly among early entrants into abnormal populations, the personality factors themselves may be so maladaptive as to create difficulty regardless of any demands made.

The present study is an attempt to evaluate some of these factors as they affect the incidence of highest points on Minnesota Multiphasic Personality Inventory profiles. Knowledge of the influence of age and sex on these distributions should also help in understanding the significance of given MMPI profile patterns.

Procedure

All of the cases in *An Atlas for the Clinical Use of the MMPI* (4) were considered, and all cases without a codable high point or with any suspicion of invalidity in the profile were eliminated (profile code followed by *X*). Only the first profile code obtained in each case was considered, regardless of how many profiles or how many admissions were recorded. The selection yielded 871 cases of whom 416 were males and 455 females. The sex distribution was then compared with the original sex distribution of 458 males and 510 females in the Atlas sample by chi square, and a chi square of 0.23 was obtained between them, which suggests that they are not significantly different.

The profile codes were divided according to which scale was highest in the profile, and the proportion of the sample accounted for by each such peak was determined. The cases were further divided by age in decades, and the proportion of cases falling in each decade was determined. Because of the small number of cases 60 years old or older, they were combined with the group which fell between 50 and 60 years of age. The expected incidences of profile peaks at given age levels and for males and females were calculated and compared with the actual incidences by chi square. The interactions between sex, age, and profile peak incidence were determined, and further analyses were made to determine at which ages significant relationships occurred and on which scales they took place. Individual incidences were examined to determine the nature and extent of the relationships with age and sex.

Results

The division of the Atlas sample by profile peaks is summarized in Table 1. *D* seems to be the scale on which peaking occurs most frequently. This observation is in accord with Meehl's (6) that *D* is the most frequent peak among abnormals.

Table 2 shows the division of the sample by age. By far the greatest proportion of the sample come into a population of abnormals during the third and fourth decades, after which there seems to be a consistent decline.

Table 1

Incidence of Individuals with Various Profile Peaks in the Atlas Sample

Profile peak	Frequency	Proportion
Hs	77	.09
D	245	.28
Hy	93	.11
Pd	160	.18
Pa	54	.06
Pt	72	.08
Sc	86	.10
Ma	84	.10

Although it is not shown in the table, only 2% of the total sample were 60 or older, and most of these were very close to 60.

When the interaction of sex and profile peak incidence is tested by chi square, a value of 39.99 is obtained ($p < .001$). Table 3 shows that profiles with *Hs* as the highest point tend to be obtained by men ($p < .05$). Profiles with *Pt* as the highest point also tend to be obtained by men ($p < .02$). Profiles with *Hy* as the highest point tend to be obtained by women ($p < .001$). A trend also seems present for women to obtained profiles in which *Pa* is highest, but it does not attain significance.

Comparison of the obtained sex distributions for each profile peak shown in Table 3 with one another by chi square yields much the same results. Significantly more men seem to obtain profiles with *Hs* as the highest point than obtain profiles with *D*, *Hy*, or *Pa*. These differences are significant at beyond the .05, .001, and .01 levels, respectively. More men obtain profiles with *Pt* as the highest point than obtain profiles with *Hy* or *Pa*. These differences are significant at the .001 and .01 levels, respectively. Significantly more women obtained profiles with *Hy* as the highest point ($p < .001$) than obtain profiles with any other scale except *Pa* as their peak. More women obtain profiles with peaks on *Pa* than obtain profiles with peaks on *Hs*, *Pt*, *Sc*, and *Ma*. The first three comparisons are significant at the .01 level, the comparison with *Ma* is significant at the .05 level.

When the interaction between age and incidence of profile peaks is tested, a chi square of 102.71 is obtained, which is significant at the .001 level. The data in Table 4 suggest that significant age relationships occur with *Hs*, *D*, *Pd*, and *Sc*. The relationship between *Hs* and age is significant at the .01 level; the latter three are significant at the .001 level. Significant relationships between profile peaks and age are noted at every period studied, except that between 30 and 39.

Inspection of the expected and obtained incidences of profile peaks at the various age levels as shown in Table 4 suggests that *Hs* and *D* peaks are comparatively uncommon in early life but become much more common as the individuals entering the clinical population become older. The reverse seems to be true of *Pd* and *Sc*. In addition, *Pt* peaks seem to be more prominent in the group which enters into an abnormal population in the twenties.

Table 2

Incidence of Individuals of Various Age Levels in the Atlas Sample

Age level	Frequency	Proportion
10–19	135	.16
20–29	253	.29
30–39	244	.28
40–49	140	.16
50–75	99	.11

Table 3

A Comparison by Chi Square of Expected and Obtained Distributions of Men and Women Obtaining Given Highest Points on the MMPI Profile

Scale	Expected *N*		Obtained *N*		Chi square
	M	*F*	*M*	*F*	
Hs	37	40	47	30	5.20[a]
D	115	130	114	131	0.02
Hy	44	49	21	72	22.82[c]
Pd	75	85	82	78	1.23
Pa	25	29	19	35	2.68
Pt	34	38	44	28	5.57[b]
Sc	40	46	44	42	0.75
Ma	39	45	45	39	1.72

[a] p is at least .05.
[b] p is at least .02.
[c] p is at least .001.

Table 4

Expected and Obtained Incidences of MMPI Profile Peaks in Successive Decades

Peak scale	Age range									
	10–19		20–29		30–39		40–49		50–59	
	E	O	E	O	E	O	E	O	E	O
Hs	12	5[a]	23	13[a]	22	29	12	19[a]	9	11
D	39	15[d]	71	60	68	75	39	51[a]	27	44[d]
Hy	16	12	28	23	27	26	16	22	10	10
Pd	25	46[d]	45	57	44	37	25	11[c]	17	9[c]
Pa	8	7	15	14	15	22	8	6	6	5
Pt	11	7	20	31[b]	19	18	11	13	8	5
Sc	14	27[c]	25	28	24	17	14	8	10	6
Ma	14	16	25	30	24	20	14	14	10	10

[a] Difference is significant at the .05 level.
[b] Difference is significant at the .02 level.
[c] Difference is significant at the .01 level.
[d] Difference is significant at the .001 level.

Discussion

The close association between *Hy* as the highest point and being female suggests the early theories of hysteria as a feminine complaint (2, p. 370). The association of peaks on *Sc* with youth suggests the Kraepelinian formulation of dementia praecox as a disorder of the young (9, p. 21). The latter is also in accord with the association of the diagnosis of schizophrenia with youth among patients admitted to mental hospitals (7).

When the description of the scales as given in the manual (3) for this test is taken into account, men seem to predominate among those with vague physical complaints which do not seem to be obviously successful in getting them out of unacceptable situations. They also seem to include a greater number of individuals with obsessive-compulsive complaints, phobias, and excessive worry and doubt. Women seem to predominate among individuals with specific somatic complaints which may be alleviated easily but return under stress. They seem to show more pathology in the area of interpersonal relationships. It should be noted that the scales on which men peak are among the scales which Welsh (8) uses to suggest internalization, while *Hy*, on which women particularly tend to predominate, is among the scales used by him to suggest externalization. The data suggest that men may be given to internalization of their problems while women may tend to externalize them.

With regard to age, the scale description in the manual suggests that younger individuals coming into an abnormal population may be characterized by difficulty in their relations with the environment as typified by acting out against it or withdrawing from it. Older persons seem to have more problems with their own physical and mental functioning. In this connection, it is of interest to note that *Hs* and *D*, which are common in the older age group, are among the scales regarded by Welsh (8) as suggestive of internalization of problems, while *Pd*, which occurs among the young, is one of the scales used by him to suggest externalization. The significantly greater proportion of individuals with peaks on *Pt* is in accord with a previous finding by Brozek (1) that younger males tend to admit to more obsessive-compulsive symptoms than older males. Brozek, however, did not find significant differences on the *Hs* scale. It should be noted that Brozek was comparing absolute scale magnitudes, while this study deals with the relative position of the scale in the profile. Lauber and Dahlstrom (5) have pointed out that the utilization of absolute scale scores may obscure relationships which the utilization of profile codes may reveal.

When the data on age and sex are considered together, *Hs* and *Pt* which occur among

males also seem to show a relation with age. Since the scales on which women peak show no such relationship, this might suggest that men tend to enter into clinical populations in the twenties and forties, while women show no tendency to enter at any particular age.

Summary

An analysis of the distribution of the profile peaks of 871 cases taken from *An Atlas for the Clinical Use of the MMPI* with regard to the sex and age distributions of the sample suggests that men tend to obtain peaks on *Hs* and *Pt* and women on *Hy* and *Pa*. Peaks on *Pd* and *Sc* seem characteristic of a younger age group and peaks on *Hs* and *D* of an older. Some possible implications of this are noted. The scales on which women peak tend to imply a more social orientation than those on which men peak. Younger entrants into an abnormal population seem characterized by behavior patterns maladaptive to societal demands. Older entrants seem characterized by problems centering around their own physical and mental functioning.

Received June 26, 1957.

References

1. Brozek, J. Personality changes with age: An item analysis of the Minnesota Multiphasic Personality Inventory. *J. Gerontol.*, 1955, **10**, 194–206.
2. Dorcus, R. M., & Shaffer, G. W. *Textbook of abnormal psychology*. (3rd ed.) Baltimore: Williams & Wilkins, 1945.
3. Hathaway, S. R., & McKinley, J. C. *Minnesota Multiphasic Personality Inventory manual* (rev. ed.). New York: Psychological Corp., 1951.
4. Hathaway, S. R., & Meehl, P. E. *An atlas for the clinical use of the MMPI*. Minneapolis: Univer. of Minnesota Press, 1951.
5. Lauber, Margaret, & Dahlstrom, W. G. MMPI findings in the rehabilitation of delinquent girls. In S. R. Hathaway & E. D. Monachesi (Eds.), *Analyzing and predicting juvenile delinquency with the MMPI*. Minneapolis: Univer. of Minnesota Press, 1953.
6. Meehl, P. E. Profile analysis of the MMPI in differential diagnosis. *J. appl. Psychol.*, 1946, **30**, 517–524.
7. National Committee Against Mental Illness, Inc. *What are the facts about mental illness?* New York: Author, 1957.
8. Welsh, G. S. An anxiety index and an internalization ratio for the MMPI. *J. consult. Psychol.*, 1952, **16**, 65–72.
9. White, R. W. *The abnormal personality*. New York: Ronald Press, 1948.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Spearman Correlations Between MMPI Scores of College Students and Their Parents[1]

Andrew L. Sopchak[2]

Memorial Center for Cancer and Allied Diseases, New York, New York

The purpose of this study was to determine the relation between MMPI scores for college students and their parents. As part of their laboratory work in elementary psychology, students at the University of Maine were given the booklet form of the MMPI. The parent(s) of 25 male students and 25 female students volunteered to take the MMPI.

The tests were scored according to standard procedure and tests whose validity scores on *?*, *L*, and *F* were high, were discarded. The rank correlations are based on raw scores on the MMPI for students and their parents.

The mean age for the 50 students was 19.9 years. The 41 mothers had a mean age of 47.7 years, and the 35 fathers had a mean age of 50.5 years.

The mothers' education ranged from high school graduates to normal school training or college graduates. Some of the fathers had only an eighth-grade education, while others were college graduates, holders of master's degrees, or were dental and medical school graduates. The mothers were mostly housewives. The father were painters, lobster men, managers, engineers, dentists, and physicians.

Results

Tables 1 and 2 present the ranges, means, and standard deviations on the MMPI scales for the students and their parents. The scores of the students are similar to scores presented by other investigators (1, 2, 5, 6, 7, 8) on other subjects. Although the present samples are relatively small, the mean scores are about the same as those obtained by Sopchak (7) on a much larger sample from the same geographical area.

In comparison to Schmidt's (5) adult men, both of the groups of fathers are considerably different. However, Schmidt's subjects were mental hospital patients. They are also different from the original standardization group of McKinley and Hathaway (3). In comparison to Verniaud's (8) adult women, both groups of women are similar to hers except on the "neurotic triad" scales on which our women are considerably higher. The means for the fathers and mothers are very similar. The mean scores of the fathers and mothers are more nearly alike than are the scores of the subjects with their parents.

Table 3 presents the Spearman correlations between the scores for the subjects and the scores for their parents.

More significant correlations were obtained between the scores of males and fathers than between those of males and mothers. On the other hand, many more significant correlations between the scores of females and their mothers than between females and their fathers were obtained. It appears that males "identify" with their fathers more than with their mothers, while females "identify" much more with their mothers than with their fathers.

Five significant correlations were obtained between the males and their fathers, but only three between the females and their fathers. Only three significant correlations were obtained between the males and their mothers, but nine between the females and their mothers.

[1] This study was supported in part by Research Grant M-884(C) from the National Institute of Mental Health, National Institutes of Health, US Public Health Service.

[2] Formerly at the University of Maine where these data were obtained.

Table 1

Ranges, Means, and Standard Deviations in *t*-Score Units for 25 Males and Their Parents

Group		*Hs*[a]	*D*	*Hy*	*Pd*[a]	*Mf*	*Pa*	*Pt*[a]	*Sc*[a]	*Ma*[a]
Men (25)	Range	42–88	33–80	43–80	29–79	12–82	33–67	9–93	44–101	38–73
	Mean	52.7	53.4	56.0	53.9	60.0	51.5	56.7	57.6	58.4
	SD	8.5	10.7	8.9	10.4	14.6	9.1	15.8	12.0	8.8
	% above 70	4	8	4	8	24	0	8	16	16
Fathers (20)	Range	36–86	34–89	44–80	29–69	41–65	41–65	36–72	32–73	35–65
	Mean	56.9	55.6	56.5	52.0	54.1	52.1	53.0	50.2	50.5
	SD	12.6	11.4	9.9	9.6	8.9	8.9	10.0	10.9	9.9
	% above 70	15	5	10	0	0	0	5	5	5
Mothers (21)	Range	38–82	44–78	42–72	32–76	28–95	35–62	23–73	38–69	27–65
	Mean	56.3	59.5	57.0	52.3	51.8	51.0	51.5	50.4	44.9
	SD	9.8	8.1	8.7	11.5	19.4	7.5	12.9	9.0	10.5
	% above 70	9	9	9	9	19	0	9	0	0

[a] Corrected with the *K* factor.

Table 2

Ranges, Means, and Standard Deviations in *t*-Score Units for 25 Females and Their Parents

Group		*Hs*[a]	*D*	*Hy*	*Pd*[a]	*Mf*	*Pa*	*Pt*[a]	*Sc*[a]	*Ma*[a]
Women (25)	Range	42–63	32–69	42–64	30–79	32–82	35–65	22–69	40–95	30–78
	Mean	49.6	47.8	55.6	54.0	52.0	54.1	51.7	57.1	56.0
	SD	6.3	9.0	5.3	10.0	10.9	7.4	10.9	12.0	10.4
	% above 70	0	0	0	8	4	0	0	12	12
Fathers (18)	Range	43–60	36–70	45–71	39–81	45–78	42–70	30–79	41–65	26–78
	Mean	52.3	53.8	56.5	54.7	57.0	53.5	52.6	54.3	51.0
	SD	6.6	10.3	7.3	11.7	8.7	8.3	9.8	9.1	14.4
	% above 70	6	6	6	17	17	11	11	6	11
Mothers (25)	Range	39–77	36–77	38–73	36–69	28–87	38–68	32–73	41–76	33–65
	Mean	51.5	55.0	57.0	53.6	48.7	52.8	51.0	51.9	50.0
	SD	7.5	11.0	7.0	7.2	12.4	7.4	12.2	8.7	8.0
	% above 70	4	16	4	0	8	0	4	4	0

[a] Corrected with the *K* factor.

Table 3

Spearman Correlations Between Scores for Students and Scores for Their Parents on the MMPI

Groups compared	*N*	*Hs*	*D*	*Hy*	*Pd*	*Mf*	*Pa*	*Sc*	*Pt*	*Ma*
Males and fathers	20	.19	−.09	−.38*	.41*	.18	.42*	.37*	.47*	.04
Males and mothers	21	.05	−.20	.06	.18	−.15	.46*	.48*	.60**	.22
Females and fathers	18	.38	.47*	.20	.30	.13	.47*	.26	.65**	.24
Females and mothers	25	.75**	.64**	.65**	.67**	.56**	.61**	.78**	.73**	.50**
Fathers and mothers of males	16	.58*	.17	.54*	.33	−.25	.49*	.53*	.36	.55*
Fathers and mothers of females	18	−.09	.01	−.20	.28	.38	.06	−.08	.14	.28

* Significant at .05 level.
** Significant at .01 level.

There are more significant correlations between the subjects and both parents on the "psychotic triad" scales than on the "neurotic triad" scales. However, scores on all three "neurotic triad" scales are significant for the females and their mothers. A most impressive finding is the obtaining of correlations significant at the .01 level on each scale of the MMPI between the females and their mothers. An explanation for this phenomenon is not offered presently.

Five significant correlations were obtained between the fathers and mothers of males, but none between the fathers and mothers of females. The significant correlations were on the "neurotic" and "psychotic triad" scales. Why the parents of males are different from the parents of females is difficult to explain. This observation should provide a lead for future studies.

According to Sakoda *et al.* (4), out of 54 calculated statistics it is possible to obtain 6 significant at the .05 level and 3 significant at the .01 level by chance. Fourteen correlations significant at the .05 level and 11 significant at the .01 level were obtained. Almost three times as many significant correlations were obtained as could have resulted by chance. It is, therefore, very unlikely that the significant correlations obtained are due to chance factors alone.

Summary and Conclusions

Over a period of four years, the group form of the MMPI had been administered to many students in elementary psychology. The parents of 25 men and 25 women volunteered to take the MMPI. This study reports on the relationship between the MMPI scores for these students and their parents.

The results permit the following conclusions:

1. The subjects and their parents are "normal" according to the criteria of the authors of the test.

2. The scores of the fathers and mothers are different from the scores of men and women in the original standardization group.

3. The scores for fathers and mothers are more similar to each other than they are to the scores of the subjects.

4. It appears that the fathers of women are more variable than the fathers of men. The mothers of men are more variable than the mothers of women. This is an interesting and challenging observation which deserves further investigation.

5. There are more significant correlations between scores of males and fathers and scores of females and mothers than between scores of males and mothers and females and fathers.

6. Significant correlations on the "psychotic triad" scales were obtained between the male and female subjects and both parents.

7. On the "neurotic triad" scales, the largest number of significant correlations was between females and their mothers.

8. Correlations significant at the .01 level were obtained for each of the nine MMPI scales between the scores of females and their mothers. No explanation for this unusual finding was offered.

9. Scores for fathers and mothers of males correlated significantly on five MMPI scales.

10. Scores for fathers and mothers of females did not correlate significantly on any of the scales. The reason for the discrepancy in correlations between parents of males and parents of females remains unexplained.

Received June 17, 1957.

References

1. Brown, H. Similarities and differences in college populations on the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1948, **32**, 541–549.
2. Lough, Orpha M. Teachers college students and the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1946, **30**, 241–247.
3. McKinley, J. C., & Hathaway, S. R. *Manual for the Minnesota Multiphasic Personality Inventory*. New York: Psychological Corp., 1951.
4. Sakoda, J. M., Cohen, B. H., & Beall, G. Test of significance for a series of statistical tests. *Psychol. Bull.*, 1954, **51**, 172–175.
5. Schmidt, H. O. Test profiles as a diagnostic aid: The Minnesota Multiphasic Inventory. *J. appl. Psychol.*, 1945, **29**, 115–131.
6. Sopchak, A. L. Parental "identification" and "tendency toward disorders" as measured by the Minnesota Multiphasic Personality Inventory. *J. abnorm. soc. Psychol.*, 1952, **47**, 159–165.
7. Sopchak, A. L. College student norms for the Minnesota Multiphasic Personality Inventory. *J. consult. Psychol.*, 1952, **16**, 445–448.
8. Verniaud, Willie M. Occupational differences in the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1946, **30**, 604–613.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

The MMPI in Cerebral Palsy

Tom Linde and C. H. Patterson[1]
University of Illinois

The problem of whether there are personality characteristics associated with the various disabilities, physical and psychosomatic, has been the object of considerable attention (1). Among others, there has been discussion of the ulcer personality, the tuberculosis personality, and the multiple sclerotic personality. The present study approaches the study of the personality of individuals with cerebral palsy, through use of the Minnesota Multiphasic Personality Inventory. It is the opinion of the authors that personality types associated with specific disabilities do not exist. Harrower and Herrmann (5), reviewing studies of multiple sclerotic patients concluded that there is no evidence of a multiple sclerosis personality. There have been no adequate objective studies of those with cerebral palsy.[2]

Sample

It is estimated that there are 550,000 cerebral palsied people in the United States, about half of them unable to work because of the severity of the disability. It was felt that if a cerebral palsy personality existed, it would manifest itself in those most severely disabled. The sampling of handicapped or disabled persons, even though they are numerous, is exceedingly difficult. Usually, intact groups or selected individuals are most conveniently available. The present sample is not a random sample, but it is felt to be rather representative of adult, severely disabled, cerebral palsied individuals, without mental or intellectual impairment, in the midwest.

The sample includes 14 males and 19 females, ranging in age from 20 to 43 years, with a mean and median age of 24. Twenty had a speech defect; nine were confined to wheel chairs all the time and three part of the time; three used crutches, and the remaining seventeen were ambulatory. Only three had full use of both hands, the remainder having at least some use, however. The range in education was from four to sixteen years, the mean being 11.4 and median 12 years. Twelve had been educated in the public schools entirely, the remainder in orthopedic or a combination of public and orthopedic schools except for three who had home instruction.

The socioeconomic background of the subjects, as indicated by the occupations of the parents, was wide, ranging from unskilled labor to engineer, though apparently concentrating in the middle occupational levels.

Results

Males and females were not significantly different, except for the *Pt* scale where the difference was significant at the .05 level.[3] The sexes were pooled for further comparisons with other groups containing both sexes.

Table 1 contains the means and standard deviations for the present sample, together with the results of statistical comparisons with (*a*) the MMPI norm group, (*b*) Canter's (3) multiple sclerosis group, and (*c*) Shontz' (7) multiple sclerosis group. Com-

[1] From the MA thesis of the senior author, done under the direction of the junior author.

[2] For a review of studies of the personality of the cerebral palsied, see the senior author's MA thesis, entitled "Personality elements of thirty-three adults with cerebral palsy as measured on the Minnesota Multiphasic Personality Inventory," on file in the University of Illinois library.

[3] All comparisons were preceded by an *F* test for homeogeneity of variances. Where variances were significantly different, the Cochran-Cox test of the significance of differences between means was applied. An asterisk indicates instances in which this test, rather than the usual *t* test, was applied.

Table 1

Means and Standard Deviations on MMPI Scales for 33 Cerebral Palsied Adults, and Statistical Results of Comparisons with Other Groups

Scale	33 CPs Mean	Comparison with manual norms $N = 800$		Comparison with Canter's MS $N = 33$		Comparison with Schontz' MS $N = 16$	
		t	p	t	p	t	p
?	32.00						
L	51.00						
F	49.00						
K	51.80						
Hs	55.55	4.14*	.01	9.90*	.01	12.65*	.01
D	52.27	1.45	>.05	10.05*	.01	4.24*	.01
Hy	56.48	4.44	.01	8.40	.01	3.53	.01
Pd	54.36	2.45	.02	1.80	>.05	.52	>.05
Mf	54.67	2.39	.02	.33	>.05	.21	>.05
Pa	55.27	3.23	.01	.91	>.05	.83	>.05
Pt	56.36	3.48	.01	2.01*	>.05	.91	>.05
Sc	63.36	8.25	.01	.08*	>.05	1.35	>.05
Ma	60.76	4.11	.01	2.19	>.05	1.55	>.05

* Cochran-Cox test used. Others are t tests.

parisons with the latter two groups were made by reconstructing variances from *SD*s given in the publication by Canter and the extended report by Shontz.

Discussion

The comparisons with the standardization sample indicate that the sample of CPs studied was significantly higher on all but the *D* scale. The actual differences are relatively small, however, except on the *Sc* and *Ma* scales, where they reach 10 or more *T*-score points. Before ascribing importance to these results, however, two considerations are suggested. First, the scores of the CPs do not differ much, if any, from scores reported on college students (2, 4, 8), except on the *Sc* scale. It might be an interesting hypothesis worthy of exploration that college students and the handicapped, perhaps for differing reasons (in the one case superior intelligence and in the other the disability) have something in common contributing to some general elevation on the MMPI scales, whether this be a higher sensitivity or insight, or some other factor.

Secondly, an analysis of the answers to the eighteen items contributing solely to the *Sc* scale was made for 24 subjects (*S*s) for whom the MMPI answer sheets were available. Seventeen answered the item: "I have never been paralyzed or had any unusual weakness of any of the muscles" in an abnormal manner. Nine responded to the item: "Sometimes my voice leaves me or changes even though I have no cold" abnormally; eight responded to the item: "I refuse to play some games because I am not good at them" abnormally. This suggests that these, and other handicapped or disabled *S*s, raise their scores on the MMPI because of responses that are classified as abnormal but are in fact true for them. On the other hand, the disabled may, again because of factors related to their disabilities possibly, manifest evidences of actual social maladjustment or withdrawal. Ten *S*s responded to the item: "The things that some of my family have done have frightened me" abnormally; and seven responded to the item: "I have never been in love with anyone" in the deviant manner. On the other hand, only three felt that they did not make friends about as quickly as others do.

The comparisons with the multiple sclerotic samples suggest that the CPs are better adjusted, or less neurotic, than the MSs. It might well be expected that the CP, whose disability is congenital, may be better ad-

justed to his disability than the multiple sclerotic, whose disability has had a relatively recent onset, and is unpredictable in progression, leading to the possible development of neurotic concern and anxiety (**5**, p. 10). Shontz (**7**) found that his MS *Ss* were significantly higher on *Hs* and *Hy* than a chronic illness group.

While there are thus statistically significant differences between the CPs in this study and the MMPI norm group, and in other studies between *Ss* with other physical handicaps or disabilities and the test norms, the conclusion that there are specific personality types or patterns associated with particular disabilities does not appear to be warranted. Scores on tests which, like the MMPI and other questionnaire tests of personality, include many items related to physical or psychosomatic conditions are probably elevated by reason of the actual presence of physical defects or complaints. Furthermore, it appears that there is, as yet at least, no evidence of unique personality patterns or types associated with differing disabilities. It is suggested that, apart from elevations due to the actual presence of physical defects or disabilities, there may be a common tendency toward personal-social maladjustment among all handicapped individuals. It is generally agreed that a disability can lead to adjustment problems. The disability may be a source of anxiety and/or a tendency toward withdrawal from social contacts. To some extent, both of these are normal in the handicapped. Whether the deviation is greater than can be expected is a question which probably should not be answered by comparing the handicapped individual with the norms for a nonhandicapped group.

Received July 12, 1957.

References

1. Barker, R. G., Wright, Beatrice A., Meyerson, L., & Gonick, Mollie R. *Adjustment to physical handicaps and illness: A survey of the social psychology of physique and disability.* Bulletin No. 55 (rev.) New York: Social Science Research Council, 1953.
2. Brown, H. L. Similarities and differences in college populations on the MMPI. *J. appl. Psychol.*, 1948, **32**, 541–549.
3. Canter, A. H. MMPI profile in multiple sclerosis. *J. consult. Psychol.*, 1952, **15**, 353–356.
4. Goodstein, L. D. Regional responses among male college students on the MMPI. *J. consult. Psychol.*, 1951, **18**, 437–445.
5. Harrower, Molly R., & Herrmann, Rosalind. *Psychological factors in the case of patients with multiple scleroses: for use of physicians.* New York: National Multiple Sclerosis Society, 1953.
6. Hathaway, S. R., & McKinley, J. C. *Manual for the Minnesota Multiphasic Personality Inventory.* New York: Psychological Corp., 1951.
7. Shontz, F. C. MMPI responses of patients with multiple sclerosis. *J. consult. Psychol.*, 1955, **19**, 74.
8. Sopchak, A. L. College student norms for the MMPI. *J. consult. Psychol.*, 1952, **15**, 445–448.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

An Evaluation of the Comparability and Reliabilities of Two Behavior Rating Scales for Mental Patients[1]

Donald W. Stilson,[2] Donald J. Mason, Malcolm D. Gynther, and Boris Gertz

South Carolina State Hospital and the University of South Carolina

A problem frequently encountered in mental health research is that of accurately and economically evaluating the level of adjustment of hospitalized mental patients (1). Two commonly used criteria of adjustment are discharge rates and transfer to "improved" wards. These criteria suffer from several inadequacies which limit their usefulness. For example, these measures (*a*) usually are dependent upon clinical (subjective) impressions, (*b*) are often influenced by interest and/or pressure from relatives, (*c*) depend upon available facilities, (*d*) are frequently conceived as being therapeutic interventions themselves, (*f*) are seldom repeatable on the same patients, (*g*) are not applicable to certain classes of mental patients (e.g., security problems, the aged, the physically infirm), (*h*) are not consistently used from one hospital to another, and (*i*) may be related to changes in administrative policies.

Two instruments which have been designed to provide a convenient method for evaluating patient adjustment are the Multidimensional Rating Scale for Psychiatric Patients (2) and the Hospital Adjustment Scale (3). Hereafter, these scales will be designated as the MRSPP and the HAS, respectively. These instruments are scored on the basis of comparatively brief patient observation periods and are relatively free of the inadequacies described above for criteria such as discharges and transfers. The present study is aimed at determining the degree to which ratings on these two scales are comparable, as well as answering several questions regarding the test-retest reliability of the MRSPP. The administration of the MRSPP requires highly trained personnel, while the HAS can be used by ward personnel with relatively limited training. Therefore, in most instances, it would be more economical to use the HAS if the measurement of patient adjustment obtained on the two instruments are highly correlated.

The degree to which the two instruments correlate with one another is, of course, a function of the reliabilities of each of the instruments. A further purpose of the present study is to assess the reliability of these scales. Many of the practical applications of the MRSPP would require repeated measurements (of the same patient group) separated by some period of time. Such measurement would presumably be sensitive to whatever changes in patient behavior had been produced by treatment. Yet, differences in ratings might be a product of a number of factors quite unrelated to the effects of therapy. For example, the test-retest reliability of the instrument may vary as a function of the type and degree of pathology shown by the patient. A cyclic manic-depressive may exhibit far greater day-to-day variation in behavior than would a severely regressed chronic schizophrenic. Similarly, the amount of personality disorganization demonstrated by any given schizophrenic might influence the reli-

[1] The authors wish to express their gratitude to Martha Bradley, Director of Nursing Education at South Carolina State Hospital, for her numerous contributions to the execution of this research. Thanks are also extended to Ruth A. Gynther, Elmore A. Martin, Jr., and Charles Presher who aided in rating patients and assembling data.

[2] Now at the University of Colorado.

ability with which he could be rated. Another factor in evaluating the test-retest reliability of an instrument is whether or not the same rater makes both ratings. Aside from the obvious chance of contaminated ratings due to a rater's recall of his previous scorings, better rapport might be established as a function of the patient and rater having discussed the same materials earlier.

To summarize, the present study attempts to assess the correlation between HAS scores and morbidity scores of the MRSPP and to evaluate the test-retest reliability of the MRSPP as a function of: (*a*) whether same or different raters make the pre- and post-ratings and (*b*) the degree of morbidity exhibited by the patients as measured by HAS scores.

Procedure

Subjects

Thirty-six *S*s were drawn from a group of 240 chronic psychotic, female patients 23 to 55 years of age. All had been diagnosed "schizophrenic" by the psychiatric staff. The group from which the experimental sample was drawn was housed in two buildings, each of which consisted of four wards. Patients were assigned to wards according to some relatively crude evaluation of morbidity. Patients on Ward A, for example, were generally less aggressive, more communicative, and showed fewer signs of severe regression than those on Ward D. Thus, a stratified population was defined with HAS total scores, building, and wards forming the three strata.

Use of the HAS

The HAS total scores obtained by student nurses approximately one month before the beginning of the present study were available for each of these patients. Before their attempt to rate members of the experimental sample, the 57 student nurses were given three two-hour training sessions and two practice ratings with patients not in the experimental sample. In the initial training, the meaning of each item on the HAS was discussed and clarified. The student nurses were then divided into groups of from three to ten members. Each of the members of a group rated the same patient whom she had observed for at least one month. In the training session following each practice rating, items showing less than 75% agreement between raters were discussed, and rules regarding the use of these items were prescribed. The distribution of the HAS total scores on the experimental sample was then divided into equal thirds. The first group consisted of scores of 33 or less; the second, of scores from 34 to 74, and the third of scores of 75 and over.

Use of the MRSPP[3]

Seven psychologists served as raters on the MRSPP. Two had the M.A. degree with three and seven years of clinical experience. The remaining five raters had doctoral degrees and have had from one to twelve years of clinical experience. During a training period, each rater administered the scale at least three times, with two or more raters evaluating the same patient. Following each rating, the results for the different raters were compared item by item, and those items which most frequently yielded disagreement were discussed by all seven raters. Specific "conventions" were adopted concerning those items which most frequently showed disagreement in an attempt to reduce interrater variability.

The interview procedure was standardized with respect to (*a*) the opening comments of the rater, (*b*) the use of certain questions which it was believed might elicit responses which were relevant to items on the MRSPP, and (*c*) questions regarding specific information (e.g., orientation) which was necessary to complete the ratings.

Patients were interviewed about one-half hour for each rating. The interviews were conducted in private rooms on the ward. An aide or nurse was then interviewed in order to obtain the information required to complete the last 22 items of the MRSPP. The same aide or nurse was always interviewed in obtaining test and retest information for a given patient.

[3] The rating conventions and interviewing guide have been deposited with the American Documentation Institute. Order Document No. 5538, remitting $1.25 for microfilm or $1.25 for photocopies.

Experimental Design

The basic design of the experiment was a $2 \times 2 \times 3$ factorial in which the variables were: (*a*) same or different rater on test and retest, (*b*) test versus retest scores, and (*c*) low, medium, and high HAS scores. Six patients were assigned to each of the 12 cells on the basis of the population stratification mentioned earlier. Each patient appeared in both test and retest categories. The samples were drawn randomly within each combination of the population strata, with sample probability proportional to the population frequency.

No rater was paired with another rater more than once under the "different rater" condition, and each rater obtained approximately half his ratings under the "test" condition. Insofar as possible, each rater obtained an equal number of ratings from each ward and each building. "Test" and "retest" ratings were obtained three weeks apart.

Results and Discussion

The data obtained in this study are summarized in Table 1. Inspection of this table reveals that the only significant main effect was HAS score ($p < .01$). This indicates a statistically significant relationship between the MRSPP morbidity scores and HAS total scores. The Pearson product-moment correlation between these two variables is $-.57$. The HAS-MRSPP correlation corrected for attenuation was $-.72$. The negative value refers to a positive relationship as high HAS scores, and low MRSPP scores are interpreted as signifying good adjustment. When test and retest scores were obtained two weeks apart using student nurse raters, the reliability of the HAS was estimated to be .79. With a three-week period intervening, the MRSPP morbidity scores yielded an .80 reliability coefficient. All three of these correlations are significant beyond the .01 level of confidence. The fact that the Rater and Test-Retest variables were not significant suggests that: (*a*) it makes no difference whether same or different raters obtain pre- and postmeasures, and (*b*) the MRSPP is sufficiently reliable to warrant its use in evaluation studies requiring repeated measurements on the same patients.

Table 1

Summary of the Analysis of Variance

Source	*df*	m.s.	*F*
HAS	2	3708.88	13.90
Test-Retest	1	4.50	0.02
Rater	1	58.67	0.22
HAS by Test-Retest	2	19.73	0.07
HAS by Rater	2	304.55	1.14
Test-Retest by Rater	1	14.22	0.05
Test-Retest by HAS by Rater	2	43.87	0.16
Within cells	60	266.82	
Total	71	341.42	

Further inspection of Table 1 indicates that none of the interactions approach significance at the .05 level. The fact that the HAS by Test-Retest interaction is not significant suggests that reliability of group means is not a function of level of morbidity as measured by the HAS. The lack of significance for the HAS by Rater interaction casts doubt on the hypothesis that repeated interviewing of the same patient has an effect on the degree of morbidity inferred by the examiner. The finding of nonsignificance for the Test-Retest by Rater interaction demonstrates that reliability is not contingent on whether the same rater obtains both test and retest ratings.

Summary

This experiment was devised to evaluate the utility of two scales widely used to rate patient morbidity. A further consideration was the reliability of each of these scales. In addition, the test-retest reliability of the MRSPP was studied as a function of degree of patient morbidity and of same or different rater on test and retest.

Estimates of test-retest reliability of the HAS and MRSPP morbidity scores were .79 and .80, respectively. The correlation between these two scales was found to be $-.57$. Further findings with the MRSPP were: (*a*) same or different raters may be used for obtaining "pre-" and "post" measures, (*b*) reli-

ability of group means is unrelated to level of morbidity, and (*c*) rating a patient more than once has no effect on the degree of morbidity inferred.

Both the MRSPP and the HAS were found to be reliable instruments for repeated ratings of psychiatric patients. For those interested in economy, the feasibility of using nonprofessional raters with the HAS is of considerable interest.

Received June 20, 1957.

References

1. Lorr, M. Rating scales and check-lists for the evaluation of psychopathology. *Psychol. Bull.*, 1954, **51**, 119–127.
2. Lorr, M., Jenkins, R. L., & Holsopple, J. Q. Multidimensional scale for rating psychiatric patients, hospital form. *U. S. Vet. Adm. tech. Bull.*, TB 10–507, Nov. 16, 1953.
3. McReynolds, P., Ballachey, E. L., & Ferguson, J. T. Development and evaluation of a behavior scale for appraising the adjustment of hospitalized patients. *Amer. Psychologist*, 1952, **7**, 340. (Abstract)

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

The Relation of Difficulty in Drawing to Ratings of Adjustment Based on Human Figure Drawings

Marvin J. Feldman and Raymond G. Hunt
University of Buffalo

The clinical use of the human figure drawing proceeds from the assumption that a subject's style of expression reflects important facets of his personality. More specifically, it is assumed that those aspects of the self subsumed under the concept of "body image" will find direct expression in the drawing of the human figure. This point of view has been elucidated most fully by Machover (1). Her elaborations of the meaning of figure drawings seem to stem in part from empirical clinical findings. In large part, however, she implies that the phantasy life of the subject is communicated by a symbolic body language. Her interpretations often parallel psychoanalytic notions about symbolism, e.g., a concave mouth reveals concern with dependency needs.

Personality may be validly revealed in many aspects by the individual's style of expression. Nevertheless, the body language conveyed by figure drawings may not have such an exact syntax or logic as implied by Machover. Further, an interpreter may be led astray by influences other than symbolic representations. Whitmyre (3) has shown rather conclusively that the over-all aesthetic value of the drawing is highly related to clinical ratings of adjustment. In his study, ratings of personal adjustment correlated .88 and .86 with independent ratings of artistic excellence for the same drawings. It might be argued that any disturbance reflecting an inadequate body image would result in a departure from a smooth, regular drawing; hence, the over-all result would be displeasing. There is no good a priori reason for believing, however, that a self depreciatory body image would lead to irregularities in drawing in just those parts of the human figure which are most difficult to draw. The present study investigates possible structural factors in drawing apart from personality needs or traits which might unwittingly influence interpretations of figure drawings.

Procedure

As part of a larger project, 65 students, 39 men and 26 women, in a lower division course in psychology were asked to draw a human figure without clothes. The drawings were obtained in a regular classroom session. Although not specifically instructed to do so, all subjects drew their own sex. Three clinicians versed in the use of figure drawings independently rated 25 body parts on each drawing for specific signs of disturbance. The ratings were simple presence or absence of disturbance for each body part. To promote uniformity of ratings, the clinical judges were given brief descriptions of the signs of disturbance associated with each body part according to Machover (1), e.g., for nose—overly long or large, tiny "button nose," heavily shaded or cut off, heavy reinforcing of lines, nostrils present.

Two art instructors in figure drawing [1] who knew nothing about the study were asked to rate the same 25 body parts in terms of the ease or difficulty in learning to draw each body part. These ratings were done on a five-point scale in which 1 indicated the easiest and 5 the most difficult part to learn to draw.

This study has no direct bearing on the validity of the symbolic meaning of signs on

[1] The authors wish to thank Anne Garson and Alexander Russo of the Albright Art School of the University of Buffalo and Saul Siegel and Herbert Sanderson who served as judges in this study.

Table 1

Interjudge Reliability for Ratings of Disturbance

Judge	r
A vs. B	.77
A vs. C	.74
B vs. C	.86

figure drawings. It would seem to furnish some evidence of a possible source of error in figure drawing interpretation.

Results

The interrater reliability of the art instructors' ratings was only .50. Inspection of the data, however, shows that they disagreed with one another by more than one step on only three items. Hence, it seemed reasonable to combine their judgments into an average rating for each body part.

The interrater reliability of the clinical judges is shown in Table 1. These correlations indicate how well the judges agreed on which body parts (not which subjects) were most often rated as manifesting disturbance. It can be seen that the clinical judges agreed fairly well in checking signs of disturbance on the various body parts of the figure drawings. Inspection of the data, however, shows that one of the judges consistently checked more signs of disturbance in rating the drawings. Hence, instead of using a raw score average of the number of times each body part was checked, it seemed advisable to convert each clinical judge's ratings to a set of standard or z scores. The three sets of z scores were then averaged.

The main comparison was then made by correlating the average ratings of the two art instructors with the average z scores of the three clinical judges. The correlation was $-.53$, which is significantly different from zero at the .01 level. This negative correlation indicates that those body parts which are most difficult to draw tend also to be rated more frequently as manifesting signs of emotional disturbance.

Table 2 presents the various body parts ranked according to artists' ratings of drawing difficulty and frequency with which they are selected by clinicians as indicating emotional disturbance. According to the body-image theory underlying Machover's rationale, one would expect certain parts of the body to be selected by clinicians more frequently than others as reflecting emotional disturbance. It follows from Schilder's discussion (2) that those parts of the body associated with erotogenic zones (e.g., extensions, openings, genitalia, etc.) will be more highly cathected and hence more likely to provide a locus of disturbance in figure drawings. Reference to Table 2 reveals that this hypothesis is only partially fulfilled. Certain supposedly (theoretically) highly cathected body parts such as breasts, nose, eyes, etc. are conspicuous by their absence among the parts most frequently selected by clinicians.

Moreover, it will be noted that those body parts which do conform to expectancy by being frequently selected by clinicians tend

Table 2

Ranks of 25 Body Parts for Artists' Ratings of Difficulty of Drawing and for Frequency of Selection by Clinicians as Indicants of Emotional Disturbance

Body part	Disturbance frequency rank	Drawing difficulty rank
Hands	5	1.5
Fingers	2	3.5
Distribution of hair over head	20	7
Distribution of hair over body	16	13.5
Nose	10	13.5
Waist	19	23.5
Ears	24	13.5
Chin	18	13.5
Body build	12	20
Height	25	23.5
Trunk	15	13.5
Joints	21	7
Shoulders	9	13.5
Arms	3	7
Chest	17	20
Breasts	14	20
Hips	11	13.5
Legs	6	7
Feet	4	1.5
Posture	13	25
Weight	22.5	20
Sex organs	1	20
Mouth	7	3.5
Eyes	8	7
Eyebrows	22.5	13.5

also to be body parts rated by artists as difficult to draw (e.g., feet, hands, etc.). Similarly, those body parts which fail to conform to expectancy, by being relatively infrequently selected by clinicians, tend to be parts rated relatively easy to draw (e.g., nose, breasts, etc.).

The only important exception to this is the clinician's rating of sex organs as the most frequent source of signs of disturbance. Sex organs were rated as easy to draw. Thus, it seems that with respect to sex organs, clinician's judgments are determined by factors other than artistic excellence. On the other hand, in terms of body-image theory, it is difficult to account for the absence of such body parts as breasts and eyes from those most frequently selected by clinicians, except as a function of some "artistic" variable.

Discussion

As was indicated above, the present study makes no necessary presumptions concerning the validity of symbolic meanings of signs on the figure drawing. Rather, the focus is upon a possible source of error confounding clinical interpretations from such drawings.

It seems quite evident from the data obtained in the present investigation and from related studies by Whitmyre (3) and Woods and Cook (4) that a considerable portion of the variance in figure drawing interpretation may be attributed to structural aspects of the drawing as distinct from symbolic aspects presumed to reflect personality traits or dimensions. As Table 2 indicates, there is considerable overlap between those parts of the figure most frequently selected by clinicians as containing indications of emotional disturbance and those parts rated most difficult to draw by figure drawing instructors (artists). However, the fact that this overlap is something less than complete (e.g., sex organs frequently selected by clinicians but rated relatively easy to render by artists) suggests that factors other than drawing skill are also influencing the clinicians' judgment.

Yet the fact remains that clinicians are evidently strongly influenced by the "artistic" dimension, both in their over-all evaluations of the drawing and with regard to specific areas of the drawing upon which they base judgments. Consequently, the variance attributed to this source would have to be somehow separated out from other sources of variance if the figure drawing method is to qualify for serious consideration as an effective general diagnostic instrument.

Summary

In an attempt to investigate possible structural factors in drawing, apart from personality needs or traits which might unwittingly influence clinical interpretations of human figure drawings, unclothed human figure drawings were collected from 65 undergraduate college students. Three clinicians rated 25 body parts on each drawing for specific signs of emotional disturbance. Two art instructors then rated the same 25 body parts along a five-point scale of difficulty in drawing.

A correlation of $-.53$ was obtained between average ratings of clinicians and artists for the 25 body parts. This correlation indicates that those body parts which are most difficult to draw tend also to be selected more frequently by clinicians as manifesting signs of emotional disturbance.

The implications of this and other similar findings for the use of figure drawings in personality diagnosis are discussed.

Received June 21, 1957.

References

1. Machover, K. *Personality projection in the drawing of the human figure.* Springfield, Ill.: Charles C Thomas, 1948.
2. Schilder, P. *The image and appearance of the human body.* Psyche Monographs No. 4; London: Kegan Paul, Trench, Trubner, 1935.
3. Whitmyre, J. W. The significance of artistic excellence in the judgment of adjustment inferred from human figure drawings. *J. consult. Psychol.*, 1953, **17**, 421–424.
4. Woods, W. A., & Cook, W. E. Proficiency in drawing and placement of hands in drawings of the human figure. *J. consult. Psychol.*, 1954, **18**, 119–121.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

On the Relationship Between Manifest Anxiety and Intelligence[1]

Charles D. Spielberger
Duke University

The Taylor Manifest Anxiety Scale (A scale) has been utilized extensively as a means of selecting subjects who are assumed to differ systematically in motivational level (4, 5, 19). Subjects (*Ss*) who scored high on the A scale have been found to perform relatively better on simple tasks (13, 20) and relatively more poorly on complex tasks (6, 13, 15, 21) than *Ss* who scored low. However, negative correlations between A-scale scores and various measures of intelligence have been reported in several recent studies (1, 3, 11). Kerrick (9) found that the A scale was significantly negatively correlated with six different measures of intelligence or aptitude. Grice (8), moreover, has demonstrated that inferior performance on a complex task by high anxiety *Ss* could be attributed to intellectual differences as readily as to differences in motivation. These findings have raised the question of the validity of the A scale as an *uncontaminated* measure of motivation.

On the other hand, Farber and Spence state that they have "been unable, over a period of years, to find any relation between the Anxiety Scale scores of college students and conventional measures of intellectual ability such as entrance-examination scores and grade point averages" (7, p. 10). Klugh and Bendig (10), Dana (2), Mayzner *et al.* (12), and Sarason (16) have also reported essentially zero correlations between the A scale and various measures of intelligence. Schulz and Calvin (17), who also failed to find a relationship between the A scale and intelligence, suggested that the contradictory findings cited above might be attributed to variations in selection procedures at different institutions. The present study is concerned with factors related to the selection of *Ss* which may affect the relation between the A scale and intelligence in college students.

Method

The A scale was administered to students in the introductory psychology course at Duke University for six consecutive semesters from the fall of 1954 through the spring of 1957. A form of the test consisting of the 50 anxiety items plus the 15 L-scale items of the MMPI was utilized. McCreary and Bendig (14) have demonstrated that the A-scale items alone correlated .96 with the same items embedded in the 225-item "Biographical Inventory" employed by Taylor (18). The ACE Psychological Examination for College Freshmen (1949 edition) is routinely administered to all entering freshmen at Duke University as part of a battery of placement tests. Scores were available on both scales for 1142 *Ss*. The data were analyzed separately for each sex for the total sample and for semester subsamples.

Results and Discussion

The means and *SDs* of the A-scale and ACE distributions and the Pearson product-moment correlations for the total sample and the various subsamples are summarized in Table 1.[2] The correlations between A scale and ACE for the total sample and for total males and total females were not significantly

[1] A paper based on part of the data of this investigation was presented at the meetings of the Midwestern Psychological Association in Chicago, May 3, 1957.

[2] The author acknowledges with appreciation the statistical and clerical assistance rendered by Ann L. Wescott and Sandra Eisdorfer.

Table 1

Means, Standard Deviations, and Correlations for A-Scale and ACE Distributions

Group	Males						Females					
	N	Mean ACE	ACE *SD*	Mean MAS	MAS *SD*	Pearson *r*	*N*	Mean ACE	ACE *SD*	Mean MAS	MAS *SD*	Pearson *r*
Spring, 1955	72	111.4	21.5	13.1	7.9	−.34**	18	128.7	21.4	10.7	5.5	.62*
Spring, 1956	79	116.8	21.1	11.3	6.3	−.21	44	124.5	16.9	11.9	7.1	−.02
Spring, 1957	101	120.5	20.9	15.5	7.2	−.16	38	134.3	17.8	14.4	6.2	−.17
Fall, 1954	140	120.3	32.6	12.4	7.8	−.13	119	128.2	18.3	13.4	5.9	.01
Fall, 1955	122	124.2	20.7	12.0	6.4	−.08	130	131.5	39.5	13.3	6.4	.09
Fall, 1956	160	124.6	18.2	14.8	7.0	.04	119	131.1	17.0	15.4	8.2	−.19*
Spring	252	116.8	21.9	13.6	7.2	−.19**	100	129.0	18.6	12.6	6.7	.08
Fall	422	123.0	24.6	13.2	7.2	.01	368	130.3	31.8	14.0	6.9	.00
Total	674	120.7	23.8	13.4	7.2	−.06	468	130.0	25.9	13.7	6.9	.01
Total sample	1142	124.5	25.1	13.5	7.1	−.02						

* $p < .05$.
** $p < .01$.

different from zero. Although two of the correlations for the females were significant, intensive analysis of the distributions of both variables for the individual and pooled samples failed to suggest any systematic interpretation of these findings.[3] Therefore, the remainder of this report is concerned only with male *Ss*.

Correlations between A-scale and ACE scores for semester subsamples for the males ranged from − .34 to + .04. Only one of these *r*'s was significantly different from zero (two-tailed tests), but the magnitudes of the obtained correlations were found to be related to the mean ACE scores of the subsamples. The relationship between mean ACE and A-scale–ACE correlation is depicted in Fig. 1A. As the mean ACE increased, the extent of the negative correlation between the A scale and the ACE decreased. The rank difference correlation for this relationship was .94.

Since the *SDs* for the various subsamples of males were approximately equal, the fact that the negative correlations between the A scale and the ACE *decreased* and approached zero as the subsample ACE means *increased* suggested that the obtained correlations were due primarily to the presence of *Ss* scoring in the lower ranges of the ACE distribution. The distributions of ACE scores for male semester subsamples were compared with the 1949 national ACE norms in Table 2. Chi-square tests of goodness of fit of the obtained distributions to the ACE norms were computed. Rank difference correlations between the percentage of *Ss* with ACE scores below the 20th and 40th percentiles of the ACE norms in a given subsample and the A-scale–ACE correlation for that subsample were − .94 and − 1.00. There was also a perfect positive rank difference correlation between the magnitude of deviation of a given subsample from the ACE

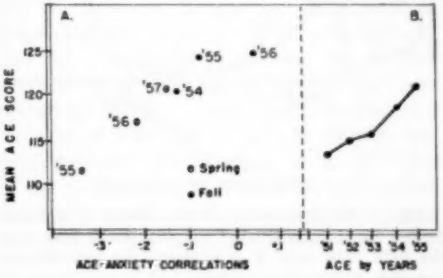

Fig. 1. *A*. Mean ACE scores and A-scale—ACE correlations for each of the six semester subsamples for males. *B*. Mean ACE scores for male freshmen by year of entry.

[3] The lower end of the range of ACE scores for the female sample was severely curtailed. Only .7% of the female *Ss* scored below the 20th percentile of the national ACE norms and only 6.8% scored below the 40th percentile.

Table 2

Distribution of ACE Scores for Males in Semester Subsamples

Group	Below 88	89–102	103–114	115–128	Above 129	Chi square
ACE norms	20%	20%	20%	20%	20%	—
Spring, 1955	14%	25%	14%	24%	24%	4.5
Spring, 1956	9%	18%	20%	19%	34%	16.0*
Spring, 1957	4%	12%	24%	27%	33%	28.0**
Fall, 1954	7%	8%	21%	26%	38%	47.0**
Fall, 1955	3%	10%	14%	32%	41%	65.6**
Fall, 1956	2%	10%	15%	34%	39%	81.0**

* $p < .01$.
** $p < .001$.

norms as measured by chi square and the A-scale–ACE correlation.

Two types of selection factors seemed to operate simultaneously to affect the number of *Ss* falling in the lower ranges of ACE scores for the male subsamples in this study. The trend of the ACE subsample means, which is depicted in Fig. 1A, indicated (*a*) that male *Ss* who took the introductory psychology course in the fall tended to be brighter than those who took it in the spring (fall mean ACE 123.0; spring mean ACE 116.8; $F = 11.1$, $p < .001$), and (*b*) that mean ACE scores increased in successive spring subsamples and successive fall subsamples as a function of time. A number of poorly understood intra-institutional selection factors appeared to be responsible for the fact that brighter students tended to take the introductory course in the fall.[4] There were no differences, however, in either the means or the variances of the distributions of A-scale scores for fall and spring. Therefore, the differences *between* the fall and spring A-scale–ACE correlations seemed to result from the differences in the distributions of ACE scores.

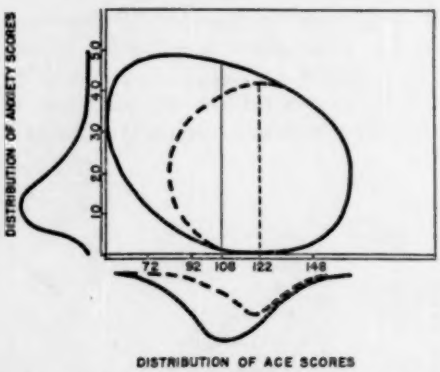

Fig. 2. An explanation of the effects of selection factors on the relationship between the A scale and the ACE. The solid lines represent the hypothetical sampling distributions of ACE and A-scale scores in an unselected population of college males. The dotted line represents the effects of selection factors which curtailed the range of the ACE distributions for this study (see Table 1 and Table 2).

The trends *within* the fall and the spring subsamples seemed to reflect the results of selection procedures which have systematically reduced the number of male students with low intelligence who have been admitted to Duke University. The mean ACE scores for entering male freshmen for the period of time affecting this study (fall, 1951—fall, 1955)[5] are shown in Fig. 1B. The trend of

[4] The introductory psychology course at Duke University is a one-semester course, and staff members who have been associated with this course over a number of years have opined that in the spring this course attracts a greater number of students who have failed the first semester of a two-semester introductory course in another area. Also, the free elective system at the university is such that students who are out of step or more poorly organized are likely to take the introductory course in the spring.

[5] The data for entering freshmen were from those students who first entered Trinity College (male undergraduate arts and sciences) during the fall semester. Data were unavailable for male freshmen who entered during the spring and summer, but this amounted to only 7% of all Trinity College freshmen.

increasing mean intelligence in successive classes of entering freshmen over this five-year period was highly significant (22).

An hypothetical explanation of these data is depicted graphically in Fig. 2, in which a small negative correlation between the A scale and the ACE is posited for male *Ss*. Where a sample contains a sizable number of cases in the lower ranges of ACE, the negative correlation between ACE and anxiety is more pronounced; where there are few cases with low ACE scores, the correlation approaches zero. Due to the operation of university selection procedures and intra-institutional selection factors in this study, the range of ACE scores is differentially curtailed at the lower end, resulting in the obtained correlations which varied as a function of the semester and year in which the introductory course was taken.

The extent to which this interpretation can be generalized to other populations is difficult to evaluate. Most studies of the relationship between anxiety and intelligence have not reported separate correlations for each sex and have not reported data on their intelligence score distributions. The above interpretation is consistent, however, with the significant negative correlations between the A scale and measures of intelligence obtained by both Grice (8) and Kerrick (9). These experimenters utilized male samples (airmen basic trainees) who would presumably be heterogeneous in intelligence with a sizable percentage of their *Ss* falling in the lower intelligence range. Matarazzo and his associates (11) found a significant negative correlation between the A scale and the ACE ($r = -.25$) in a group of male college sophomores. This correlation appears to be somewhat inconsistent with the suggested interpretation of our findings in that the mean ACE score for Matarazzo's sample is very high (estimated by the present investigator from their Fig. 2 to be approximately 134; 87th percentile on 1949 ACE norms). Although data on the ACE distribution in the Matarazzo study is not specifically reported, the range of ACE scores appeared to be moderately heterogeneous and might include a sufficient number of cases with low ACE scores to account for the obtained correlation. In this same study, the authors failed to find any relationship between the A scale and either the CVS (an abbreviated version of the Wechsler Bellevue) or grade-point averages which points up the importance of the particular measure of intelligence used. Calvin *et al.* (1) found no relationship between A-scale scores and IQ as measured by the Wechsler Bellevue Intelligence Test in a group of college *Ss* (21 males; 15 females) whose IQ range was 102–141 (mean 122.75, *SD* 9.18). When they added a second group of 15 *Ss* (12 males; 3 females) with lower scores (range 92–129, mean 108.3, *SD* 10.41), the combined groups yielded highly significant negative correlations between the A scale and IQ. The emergence of significant negative correlations as a function of including a number of male *Ss* with low intelligence is consistent with the interpretation of the data of the present study. Calvin *et al.* point out, however, that low IQ in their second group was confounded with scholastic difficulty in that *Ss* in this group were obtained from a how-to-study course offered for students having scholastic difficulty.

While Kerrick's statement that it is "virtually impossible to select extreme subjects on the Taylor scale who are equated in intelligence" is obviously not consistent with the data, the implication of the studies cited above and the findings of the present study are that a small negative correlation between anxiety and intelligence may be found in male college *Ss* if one samples a wide range of intelligence and if the sample contains a sizable proportion of *Ss* from the lower part of the range. However, this correlation may be obscured by variations in selection procedures which curtail the range of intelligence. It should be noted that the present study offers no explanation for a relationship between the A scale and the ACE, nor does it answer questions pertaining to the validity of the A scale as a measure of motivation. It does suggest that experimenters selecting male *Ss* on the basis of the A scale should control for intelligence, especially when studying heterogeneous populations with low mean intelligence.

Summary

Scores on the Taylor Manifest Anxiety Scale and the ACE Psychological Examination for College Freshmen were obtained for

six consecutive semesters from a total of 1142 students enrolled in introductory psychology courses. The correlations between the A scale and the ACE for the total sample, total males, and total females were not significantly different from zero. Although only one of six semester correlations for the males was significant, the magnitudes of these correlations were highly related to the mean ACE scores for these subsamples. Where a subsample contained a sizable proportion of *Ss* with low ACE scores, the negative correlation between the A scale and the ACE was more pronounced; where there were few *Ss* with low ACE scores, the correlation approached zero. It was suggested that a small negative correlation between the A scale and intelligence may be found in college males if one samples a wide range of intelligence and if the sample contains a number of *Ss* from the lower part of the range. However, this correlation may be obscured by variations in selection procedures which curtail the range of intelligence.

Received June 18, 1957.

References

1. Calvin, A. D., Koons, P. B., Bingham, J. L., & Fink, H. H. A further investigation of the relationship between manifest anxiety and intelligence. *J. consult. Psychol.*, 1955, **19**, 280–282.
2. Dana, R. H. Manifest anxiety, intelligence, and psychopathology. *J. consult. Psychol.*, 1957, **21**, 38–40.
3. Duvall, Mary I. The effect of manifest anxiety upon incidental and new task learning. Unpublished doctoral dissertation, George Peabody College, 1956.
4. Farber, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln: Nebraska Univer. Press, 1954.
5. Farber, I. E. The role of motivation in verbal learning and performance. *Psychol. Bull.*, 1955, **52**, 311–327.
6. Farber, I. E., & Spence, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.
7. Farber, I. E., & Spence, K. W. Main and interactive effects of several variables on reaction time. *U. S. Naval Res. Lab.*, Mar. 1, 1955, Tech. Rep. No. 3. (Contract N9 onr-93802, State Univer. of Iowa.)
8. Grice, G. R. Discrimination reaction time as a function of anxiety and intelligence. *J. abnorm. soc. Psychol.*, 1955, **50**, 71–74.
9. Kerrick, Jean S. Some correlates of Taylor Manifest Anxiety Scale. *J. abnorm. soc. Psychol.*, 1955, **50**, 75–77.
10. Klugh, H. E., & Bendig, A. W. The Manifest Anxiety and ACE scales and college achievement. *J. consult. Psychol.*, 1955, **19**, 487.
11. Matarazzo, J. D., Ulett, G. A., Guze, S. B., & Saslow, G. The relationship between anxiety level and several measures of intelligence. *J. consult. Psychol.*, 1954, **18**, 201–205.
12. Mayzner, M. S., Sersen, E., & Tresselt, M. E. The Taylor Manifest Anxiety Scale and intelligence. *J. consult. Psychol.*, 1955, **19**, 401–404.
13. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
14. McCreary, J. B., & Bendig, A. W. Comparison of two forms of the Manifest Anxiety Scale. *J. consult. Psychol.*, 1954, **18**, 206.
15. Ramond, C. Anxiety and task as determiners of verbal performance. *J. exp. Psychol.*, 1953, **46**, 120–124.
16. Sarason, I. G. The relationship of anxiety and "lack of defensiveness" to intellectual performance. *J. consult. Psychol.*, 1956, **20**, 220–222.
17. Schulz, R. E., & Calvin, A. D. A failure to replicate the findings of a negative correlation between Manifest Anxiety and ACE scores. *J. consult. Psychol.*, 1955, **19**, 223–224.
18. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
19. Taylor, Janet A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303–320.
20. Taylor, Janet A., & Chapman, J. P. Paired-associate learning as related to anxiety. *Amer. J. Psychol.*, 1955, **68**, 671.
21. Taylor, Janet A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.
22. Weitz, H. Placement test performance of Duke University freshmen—a five year summary. Duke University, Bur. of Testing and Guidance, Staff Research Report No. 14, Nov. 15, 1955.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Anxiety: Unitas or Multiplex?[1, 2]

Douglas N. Jackson
Pennsylvania State University

and Robert Bloomberg
VA Hospital, St. Cloud, Minnesota

One way of determining if a construct or dimension, such as anxiety, is unitary is to examine its relationships with other constructs and with the observables which help to define the constructs. If a construct is defined directly in terms of a number of independently measurable observables, one can look for empirical relationships among the observables which should reflect something about the construct. The conclusion that a single dimension is sufficient to account for all of the variables presumed to represent a construct like anxiety requires not only that all of the variables are substantially correlated, but also that the correlations among all of the measures are due to a single common factor to which all of the measures are mutually related. The present study seeks to assess relationships among a number of purported measures of anxiety in order to determine the degree to which these reflect a single dimension.

Method

Subjects

From the admissions to a neuropsychiatric hospital, 37 male patients were selected randomly, with the exceptions that those with known organic pathology, as well as those with ages in excess of 45 years, were excluded. The age range was 19–43 with a median of 30 years. Diagnostically, the group was composed as follows: schizophrenic reaction, 22; manic-depressive, 1; psychoneurotic (including depressive) reaction, 13; alcoholism, 1.

Procedure

Four measures were administered to each *S:* (*a*) The Wechsler-Bellevue digit span score, which Rapaport (10, pp. 185–193) has put forth as a measure of anxiety without confirmation in later studies (6), and a vocabulary scale (11) to permit statistical control of verbal intelligence in assessing digit span performance; (*b*) the Taylor Manifest Anxiety Scale (13), administered as part of the Minnesota Multiphasic Personality Inventory (MMPI); (*c*) two one-minute observations of the blink rate, recorded while the *S* viewed a drawing of a triangle mounted on a vertical plane 60 inches from his line of sight; (*d*) the fingerprint stain test of palmar sweating.

The blink rate. Meyer (7) reviews the empirical evidence for considering the blink rate as an indicant of anxiety. For example, when an individual is under stress, as he is while undergoing cross-examination in a court of law, the blink rate increases (9). Meyer offers a theoretical interpretation which attributes the interaction of simultaneous responses to a convergence of impulse patterns upon the motor pathways of the central nervous system. Variations in the blink rate are attributed to excitation from these neighbor-

[1] Portions of this paper were read at the meetings of the Midwestern Psychol. Ass., St. Louis, 1956. Most of this report was written while the first author was at the Menninger Foundation as a U. S. Public Health Service Clinical Research Fellow of the National Institute of Mental Health.

[2] Published with permission of the Chief Medical Director, Department of Medicine and Surgery, Veteran's Administration, who assumes no responsibility for the opinions expressed or conclusions drawn by the authors.

ing channels. The corrected split-half reliability of the blink rate index was .91 in the present study.

The fingerprint stain test of palmar sweating. Mowrer, Light, Luria, and Zeleny (8) review the successive developments of this device. It was validated by evidence that it decreases in therapy (1, 5) and in a relaxing situation (3), while increasing under stressful conditions. Anxiety neurotics have shown greater palmar sweating than "normals" (5).

The test, described by Mowrer *et al.* (8), consists of dabbing a solution of anhydrous ferrous chloride on the index finger and allowing it to dry for one minute. The *S* is then instructed to place his index finger (at one pound pressure as measured by a scale) on a piece of Dietzgen 198 M paper treated with a five per cent solution of tannic acid. The amount of ferrous chloride deposited on the paper, and hence the density of the fingerprint, has been shown to be a linear function of the amount of sweat (12). The density of the fingerprint is determined by the amount of light passing through the print by means of a specially constructed photoelectric circuit containing a microammeter. The repeat-reliability of this photometric reading of fingerprint density was in excess of .99 in the present study. The test-retest reliability of this measure of palmar sweating has been found to be .89 after a half-hour interval (14).

Results

Product-moment correlations were computed among the four measures. Table 1 indicates that none of these correlations are significantly different from zero. When the vocabulary score is partialled out of correlations involving the digit span measure, the range is from − .17 to + .02.

Table 1

Correlations Among Anxiety Measures

Measures	Digit span	Taylor Anxiety scale	Blink rate
Taylor Anxiety scale	.02*		
Blink rate	−.08*	−.11	
Palmar sweating	−.17*	−.15	−.02

* Covariance attributable to verbal knowledge partialled out.

Discussion

If these results concerning lack of relationship between assumed measures of "anxiety" prove reliable, and an extensive factor analytic study by Holtzman and Bitterman (4) supports this finding, a careful scrutiny of the conceptualization and experimental definition of anxiety seems in order.

One possibility is to consider anxiety as a second-order factor, and consider the primary factors to be such dimensions as muscular tension, palmar sweating, etc., analogous to the situation now prevailing in the area of human abilities. This would be true only if there was some correlation between the primary factors. The evidence on this point is incomplete, but if various proposed measures of anxiety can be placed along orthogonal axes and are completely unrelated, one no longer has a meaningful single dimension or grouping of correlated dimensions of anxiety, but is left with a less parsimonious, but possibly more realistic, structure of independent dimensions of such things as muscular tension, of palmar sweating, and, among other things, a dimension of test-taking behavior, the principal factor of which may be best exemplified by the Taylor scale or the MMPI *K* scale.

The evidence reported here cannot but raise questions concerning the generality of at least some of the measures employed. Widespread and uncritical use of the Taylor scale as an indicant of anxiety, for example, may lead to misleading results and erroneous conclusions based only to a minor extent upon variance associated with "manifest anxiety" and to a greater, if unknown, extent upon such determinants as response acquiescence (2), social desirability, or other superfluous factors.

Summary

In order to gather preliminary evidence concerning the extent to which four supposed indicants of anxiety could be considered to tap the same or related dimensions in a population of male functional psychiatric hospital patients, the following tests were administered: (*a*) digit span; (*b*) Taylor Anxiety

scale; (*c*) fingerprint stain test of palmar sweating; and (*d*) blink rate.

Although the measures were acceptably reliable, their intercorrelations failed to depart significantly from zero. These results, when considered with other evidence, suggested that manifestations of anxiety are multidimensional rather than unitary and may even represent uncorrelated dimensions. This may necessitate a conceptual reinterpretation of anxiety and cautions against the uncritical acceptance of a single measure as an indicant of generalized anxiety.

Received June 10, 1957.

References

1. Bixenstine, V. E. A case study of the use of palmar sweating as a measure of psychological tension. *J. abnorm. soc. Psychol.*, 1955, **50**, 138–143.
2. Cronbach, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt*, 1950, **10**, 3–31.
3. Gladstone, R. An investigation of the relationship between palmar sweating and emotion as measured by a group test of palmar sweating. Unpublished doctoral dissertation, Univer. of Illinois, 1949.
4. Holtzman, W. H., & Bitterman, M. E. A factorial study of adjustment to stress. *J. abnorm. soc. Psychol.*, 1956, **52**, 179–185.
5. Light, B. H. Tension changes in patients undergoing psychotherapy. Unpublished doctoral dissertation, Univer. of Illinois, 1951.
6. Matarazzo, Ruth G. The relationship of manifest anxiety to Wechsler-Bellevue subtest performance. *J. consult. Psychol.*, 1955, **19**, 218.
7. Meyer, D. R. On the interaction of simultaneous responses. *Psychol. Bull.*, 1953, **50**, 204–220.
8. Mowrer, O. H., Light, B. H., Luriz, Z., & Zeleny, M. P. Tension changes in psychotherapy with special reference to resistance. In O. H. Mowrer (Ed.), *Psychotherapy: Theory and research*. New York: Ronald Press, 1953. Pp. 546–640.
9. Ponder, E., & Kennedy, W. P. On the act of blinking. *Quart. J. exp. Physiol.*, 1927, **18**, 89–110.
10. Rapaport, D. *Diagnostic psychological testing*. Chicago: Year Book Publishers, 1945.
11. Shipley, W. C. A self-administering scale for measuring intellectual impairment and deterioration. *J. Psychol.*, 1940, **9**, 371–377.
12. Silverman, J. J., & Powell, V. E. Studies on palmar sweating. *Psychosom. Med.*, 1944, **6**, 243–249.
13. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
14. Wenger, W. A., & Gilchrist, J. C. A comparison of two indices of palmar sweating. *J. exp. Psychol.*, 1948, **38**, 757–761.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Test Anxiety Questionnaires[1]

George Mandler and Judith E. Cowen[2]
Harvard University

The increased use of paper-and-pencil tests for the measurement of anxiety makes it appropriate to communicate the following additional information on one of these instruments.

The college form of the Test Anxiety Questionnaire (TAQ) was first described by Mandler and Sarason (3). A high school form has since been developed by Cowen (1). These questionnaires were designed to elicit self-ratings on items descriptive of anxiety reactions in test situations. The college form contains 39 items, of which four are fillers; the high school form contains 52 items, including four fillers.[3] Each item consists of a graphic scale with the end points marked with written description and the midpoint indicated. The subject is free to mark each individual scale anywhere along the continuum.

Scoring

Originally, the TAQ was scored by dividing each scale into 10 intervals and assigning a score of 1 to 10 to the subject for each scale. A median was then established for the distribution on each individual scale and each subject received a score of zero or one depending on whether his scale response fell below or above the sample median. Sarason and Gordon (5) have presented scoring norms based on this system.

A preferable method of scoring the TAQ is to sum the raw scores on the individual items, thus eliminating the tedious determination of medians for each item, which may vary from sample to sample. An empirical comparison of the two scoring systems on two independent samples yielded product-moment correlations of .95 ($N = 81$) and .93 ($N = 41$). Thus, the new scoring system adequately fulfills the function of the old, while at the same time saving considerable time and effort and eliminating the need for norms specific to the sample.[4]

Reliability

The test-retest reliability of the college form of the TAQ is .91 ($N = 70$), the split-half reliability is .91 ($N = 100$) (6). The split-half reliability of the high school form of the TAQ is .90 ($N = 286$).

Norms and Sex Differences

The college form of the TAQ was administered to a group of Harvard-Radcliffe sophomores and juniors ($N = 99$) and a group of Boston University freshmen and sophomores ($N = 80$).[5] The high school form was administered to a group of Watertown, Massachusetts, high school sophomores ($N = 286$). Table 1 shows means and standard deviations for these groups by educational level and by sex.

All the means on the college form are near the midpoint of the theoretically possible

[1] Parts of this research were supported by a grant from the Laboratory of Social Relations, Harvard University.

[2] Now at Department of Education, Waterbury, Connecticut.

[3] Copies of the two questionnaires and detailed scoring instructions have been deposited with the American Documentation Institute. Order Document No. 5485, remitting $1.75 for microfilm or $2.50 for photocopies.

[4] If the TAQ is considered as a composite of 35 different scales, Gulliksen's discussion (2) of the determination of composite scores is relevant. The empirical relationship found here could be predicted from his analysis, namely, that "the test mean (has) no effect whatever on the relationship between a test and the composite of which it is a part."

[5] The data were supplied by Michael Maccoby.

Table 1

Means and Standard Deviations on Two Forms of the TAQ

Form and group	*N*	Mean	*SD*
College Form			
Harvard Radcliffe sample	99	179.8	38.6
Male subjects	66	179.8	38.5
Female subjects	33	179.7	39.6
Boston University sample	80	167.1	47.8
Male subjects	48	158.7	48.4
Female subjects	32	179.7	44.5
High School Form			
Watertown sample	286	245.9	41.1
Male subjects	151	239.17	41.5
Female subjects	135	253.55	39.2

range, which is 35 to 350. The Harvard-Radcliffe sample shows a higher mean than the Boston University sample. The only significant sex difference (women scoring higher than men) appears on the Boston University sample ($p = .05$). The Watertown sample on the high school form (which is scored on a nine-point scale) also has a mean near the theoretical midpoint (range is 48 to 432) and sex differences in the same direction as the Boston University sample, although these are not significant.

Test Anxiety and Other Measures

Since test anxiety has been hypothesized to be operative in test situations, the relation between the TAQ and intelligence tests is of interest. Sarason and Mandler (6) reported a correlation of $-.209$ ($p < .05$) between the TAQ and the Henmon-Nelson Test of Mental Ability, college form. Cowen (1) reported a similar correlation between the high school TAQ and the Otis Gamma of $-.185$ ($p < .02$).[6]

Finally, the relationship between the TAQ and the widely used Taylor scale of Manifest Anxiety (7) is of interest. Raphelson (4) reported a correlation of .53 ($N = 24$). A replication on a Harvard-Radcliffe sample resulted in a correlation of .59 ($N = 35$). This consistent positive relationship between the two scales indicates some communality of traits measured, but is too low to permit either substitution of the two scales for each other or any uncritical comparison of the results obtained by the use of these instruments.

Received May 23, 1957.

References

1. Cowen, Judith E. Test anxiety in high school students and its relationship to performance on group tests. Unpublished doctoral dissertation, School of Education, Harvard Univer., 1957.
2. Gulliksen, H. *Theory of mental tests.* New York: Wiley, 1950.
3. Mandler, G., & Sarason, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, **47**, 166–173.
4. Raphelson, A. C. The relationships among imaginative, direct verbal, and physiological measures of anxiety in an achievement situation. *J. abnorm. soc. Psychol.*, 1957, **54**, 13–18.
5. Sarason, S. B., & Gordon, E. M. The test anxiety questionnaire: Scoring norms. *J. abnorm. soc. Psychol.*, 1953, **48**, 447–448.
6. Sarason, S. B., & Mandler, G. Some correlates of test anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 810–817.
7. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
8. Zweibelson, I. Test anxiety and intelligence test performance. *J. consult. Psychol.*, 1956, **20**, 479–481.

[6] Comparable correlations between a children's test anxiety questionnaire and intelligence test scores have been reported by Zweibelson (8).

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Allergy and the Need for Recognition

Norman S. Greenfield[1]
University of Wisconsin Medical School

The number of papers which have been devoted to discussions of possible relationships between emotions and allergy would seem to make any systematic review of the literature somewhat redundant. The interested reader will find reviews of this work in Leigh (5) and in Dunbar (1).

A central focus of discussion in recent years has been the formulation generally credited to French and Alexander (2) relating allergy to maternal rejection. Support for this finding, gained during the psychoanalysis of asthmatic patients was offered by Saul (10) who reported that unsatisfied longings for love contribute to allergic sensitivity. The numerous researches of Miller and Baruch (6, 7, 8) primarily utilizing material inferred from interviews and observations of play techniques, have been widely quoted and consistently indicate that ". . . maternal rejection appears to be an important factor in the personality and somatic symptom-patterns of the clinically allergic child" (6, p. 278).

More recently, Harris and Shure used the ". . . unhampered appraisal of the teacher who observes the child day after day in the natural setting of the classroom . . ." (3, p. 322) and, basing their conclusions on teachers' descriptions of childrens' behavior, reported that no discernible personality pattern differentiates asthmatic from nonasthmatic children. Additionally, these authors wisely pointed out that the range of clinical observational techniques which have been brought to bear on the attempted resolution of this problem are extremely susceptible to the error of subjective influence.

[1] The author wishes to thank Seymour B. Crepea, Associate Professor of Medicine, for his technical assistance and helpful comments during the course of this research.

It was reasoned that if maternal rejection and a subsequent unsatisfied longing for love are indeed associated with allergic sensitivity, then a hypothesis derivable from a systematic extension of this classical formulation would be that the allergic individual would have a greater need for recognition than will individuals who do not demonstrate allergic sensitivity. This need for recognition might be thought of as an attempt to gain a kind of belated mastery over the impulses which have been frustrated during an earlier period. Thus the task of the present study was to subject to systematic experimental analysis the hypothesis that the need for recognition is greater in a group of individuals receiving treatment for allergy than in a control group.

Method

The Technique

The Iflund Test (4), which is based on the system of psychological needs contained in Murray's theory of personality (9), was administered individually to each of the subjects (*Ss*). Essentially, this is a memory test which consists of 51 pictures, 17 for men, 17 for women, and 17 for both. Thus 34 pictures are used in any single test. Each of 28 pictures is designed to represent a specific need (e.g., autonomy, exposition, abasement, recognition, etc.) and they are presented serially at constant intervals to the *S*. Three "bland" pictures are shown at the beginning and three at the end of the series in order to minimize the effects of primacy and recency of exposure. Following each series of exposures the *S* is asked to describe those he remembers. The exposures are repeated until the *S* has recalled the entire series. Some are recalled after the first showing, others after several showings. The assumption is made that the

order in which the *S* recalls the pictures is related to his personality organization. Additionally, representations of needs which are being gratified are recalled earlier, and representations of ungratified needs are recalled later. It was hypothesized that since the need for recognition of the allergic individual is manifest rather than repressed, the pictorial representation of this need would be recalled earlier. Previous investigation (4) clearly demonstrated that the order of recall is determined primarily by the meaning of the pictures which is in turn related to the personality of the individual.

In the present study, the actual testing was done by a psychologically unsophisticated research assistant who was not aware of the specific hypothesis under examination. It was felt that this would serve as further insurance against the possibility of the examiner's communicating subliminal cues to the *S*.

Subjects

The experimental group was composed of 35 students, 20 males and 15 females, who were receiving regular immunologic treatment for allergic rhinitis and/or asthma at the allergy clinic of the University Student Health Department. A purely operational definition of allergy in terms of patients who had been placed on an immunologic regimen by either their own physicians or members of the clinic staff was employed. In short, allergy was defined only in terms of those individuals who were receiving treatment for allergy. No effort was made to recheck or validate the diagnosis. Under these relatively loosely defined conditions it is quite likely that some misdiagnosed individuals were included in the experimental group. However, it can readily be seen that to the extent to which this is so, the resultant error would tend to attenuate rather than confirm the hypothesis.

The technique of matched control was used with each allergic patient being paired with a nonallergic peer for sex, age, and level of education. The mean age of the group was 20.97 years with a range from 17 to 29 years. Controls were drawn from various extracurricular college groups.

Results

In order to establish a quantitative measure or a recall score, a *T* score was computed for each need picture expressing the number of showings for that picture relative to the total number of showings required. This method tends to equate the *S*s in terms of general memory ability since the actual score is a ratio. The mean *T* score for both groups was 70.07. The mean *T* score for the allergy group was 62.2, while that for the control groups was 77.9. The nonparametric sign test was employed in the analysis of the data, and this difference was found to be significant at beyond the .01 level of confidence, demonstrating that the ideational representation of the need for recognition was recalled significantly earlier among the allergic group than among the controls. No differences between the sexes were observed nor were the groups differentiated by any of the other variables which made up the total test.

Discussion

These results suggest confirmation of the thesis that, within the definition of the samples employed, the need for recognition is prominently associated with the personalities of patients who are on an immunologic regimen. The fact that the significant difference occurred in the form of early recall suggests that this need is, at some level, being satisfied, possibly by the ritual which is associated with allergy.

It should be made abundantly clear that there is no intent to imply that these data denote a causal relationship. The present study was designed to provide a methodologically rigorous test of an essentially clinical assumption, and it is felt that this task has been fulfilled. Whether the constellation of factors which are inferable from these findings are cause for, or a reflection of, the allergic condition remains a problem for future research in this area.

Summary and Conclusion

The hypothesis that there is a greater need for recognition among a group of individuals being treated for allergic sensitivity than among a control group was tested with a uni-

versity student health population. Within the definition of the method and the sample the hypothesis was confirmed. It was emphasized that no causal relationship was implied by these findings.

Received July 1, 1957.

References

1. Dunbar, H. Flanders. *Emotions and bodily changes.* (4th ed.) New York: Columbia Univer. Press, 1954.
2. French, T. M., & Alexander, F. Psychogenic factors in bronchial asthma (Part I). *Psychosom. Med. Monogr.*, 1941, No. 4.
3. Harris, M. C., & Shure, N. A study of behavior patterns in asthmatic children. *J. Allergy*, 1956, **27**, 312–323.
4. Iflund, B. Selective recall of meaningful material as relative to psychoanalytic formulations in certain psychiatric syndromes. Unpublished doctoral dissertation, Univer. of California, 1953.
5. Leigh, D. Asthma and the psychiatrist; a critical review. *Internat. Arch. Allergy*, 1953, **4**, 227–246.
6. Miller, H., & Baruch, Dorothy W. Psychosomatic studies of children with allergic manifestations. *Psychosom. Med.*, 1948, **10**, 275–278.
7. Miller, H., & Baruch, Dorothy W. Marital adjustment in the parents of allergic children. *Ann. Allergy*, 1950, **8**, 754–760.
8. Miller, H., & Baruch, Dorothy W. A study of hostility in allergic children. *Amer. J. Orthopsychiat.*, 1950, **20**, 506–519.
9. Murray, H. A. *Explorations in personality.* New York: Oxford Univer. Press, 1948.
10. Saul, L. J. Some observations on the relation of emotion and allergy. *Psychosom. Med.*, 1941, **3**, 66–71.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Chance Frequency in Matching Problems When Success or Failure is Reported After Each Matching Operation

Alan H. Roberts
Elgin State Hospital

Despite the relative importance of matching procedures in clinical investigations, the sources most often consulted by psychologists contain very little information concerning the statistical treatment of these kinds of data. Among the exceptions are an article by Dudek (1) which provides a formula and tables for determining chance probabilities when n different categories are matched with n different objects, and a discussion by Mosteller and Bush (2) concerning the statistical analyses of several matching designs.

An extremely useful matching procedure whose treatment has been neglected is one in which n different objects are to be matched with n different categories and the judge is to be informed of the success or failure of each matching operation as he proceeds. An example of this procedure might be an experiment in which clinical psychologists are

Table 1

Chance Frequencies (Roman Type) and Proportions (Bold-faced Type) When n Items are Distributed into n Categories and Success or Failure is Reported After Each Match

Items (n)	Number of correct matches (r)										Total orders ($n!$)
	1	2	3	4	5	6	7	8	9	10	
2	1 **.5000**	1 **.5000**									2
3	2 **.3333**	3 **.5000**	1 **.1667**								6
4	6 **.2500**	11 **.4583**	6 **.2500**	1 **.0417**							24
5	24 **.2000**	50 **.4167**	35 **.2917**	10 **.0833**	1 **.0083**						120
6	120 **.1667**	274 **.3806**	225 **.3125**	85 **.1181**	15 **.0208**	1 **.0014**					720
7	720 **.1429**	1764 **.3500**	1624 **.3222**	735 **.1458**	175 **.0347**	21 **.0042**	1 **.00020**				5040
8	5040 **.1250**	13068 **.3241**	13132 **.3257**	6769 **.1679**	1960 **.0486**	322 **.0080**	28 **.00069**	1 **.000025**			40320
9	40320 **.1111**	109584 **.3020**	118124 **.3255**	67284 **.1854**	22449 **.0618**	4536 **.0125**	546 **.0015**	36 **.000099**	1 **.000003**		362880
10	362880 **.1000**	1026576 **.2829**	1172700 **.3232**	723680 **.1994**	269325 **.0742**	63273 **.0174**	9450 **.0026**	870 **.00024**	45 **.000012**	1 **.0000003**	3628800

asked to match n Rorschach protocols with n different diagnostic categories. In such an experiment, the *subjective* probabilities in favor of the judge might be increased if he were told the success or failure of each match as he proceeds so that he might be given a second chance to match a given protocol if the first match were in error. The judge matches the first protocol with one of the diagnostic categories. If the match is correct, he is so told, and the protocol together with its diagnostic category is removed. If the match is incorrect, he is so told, and the protocol together with its *correct* diagnosis is removed. Either correct protocols or correct diagnoses may be removed so long as the procedure first used is continued throughout the experiment. The procedure is repeated until n matches have been made. The experimenter must then determine if the number of correct matches achieved is likely to have arisen by chance.

When a procedure of this type is followed, the tables supplied by Dudek are not applicable. The correct chance frequencies and their corresponding proportions are shown in Table 1 for $n = 2$ to 10. Where $n > 10$, the correct frequencies can be derived from this table by means of the formula

$$X_r^n = (n-1)(X_r^{n-1}) + X_{r-1}^{n-1}$$

where X_r^n is the chance frequency of obtaining r correct matches out of n pairs.

Received June 10, 1957.

References

1. Dudek, F. J. Determining "chance success" when a specific number of items are sorted into discrete categories. *J. consult. Psychol.*, 1952, **16**, 251–256.
2. Mosteller, F., & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. 1. *Theory and method*. Cambridge: Addison-Wesley, 1954. Pp. 289–334.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

An Unprofitable Application of the Semantic Differential[1,2]

Jack Block

Institute of Personality Assessment and Research, University of California, Berkeley

Since the development of the semantic differential (2), the technique has achieved a variety of research applications. Central to a number of its usages has been the assumption that in some degree, in some fashion, the technique reveals relationships among concepts of which the subject may be unaware. If not yet deemed a measure reflecting of the unconscious, a number of investigators still have suggested the technique "pulls" information related to some underlying, complexly symbolic mediating process which is a powerful determinant of behavior but which the subject is unable to report directly.

This assumption is an important one to test, for, if it is correct, we can bring the technique to bear upon a large number of substantive problems. If the assumption is in fact unwarranted, then we are engaging in an uneconomical self-deception where more direct and simpler procedures are already available. The present note reports on an empirical test of the assumption of the "depth-penetration" of the semantic differential.

A sample of 112 students in a large introductory course in psychology at the University of California served as subjects (*Ss*). Each *S* was provided a booklet and asked to describe his ideal self and his like-sexed parent by means of the semantic differential technique. The connotative scales employed for these descriptions came from an Osgood report (2), and a seven-point rating continuum was specified.

[1] This investigation, part of a larger research project, was supported by Research Grant M-1078 from the National Institute of Mental Health of the United States Public Health Service.

[2] The writer is indebted to John Enright for his help in evaluating the results of this study.

Following the semantic differential descriptions, each *S* was asked to check a list of 80 adjectives so as to describe his ideal self and then to check the adjective list so as to portray his like-sexed parent. Adjectives were to be identified as "characteristic or true," as "uncharacteristic or false," or as not especially relevant for the description.

Where the semantic differential descriptions were recorded on separate pages of the test booklet, the adjective descriptions were recorded on immediately adjacent (and immediately comparable) columns. This arrangement of material was intended to reduce the probability of direct comparison by the *S* of his semantic differential descriptions while accentuating the likelihood of comparison with the adjective list.

From these data, two sets of scores were derived. The similarity of one's ideal self to one's parent represents a kind of identification (4, 5). The similarity of the two semantic differential descriptions provides one measure of this identification and was quantified by computing the "distance" (1, 3) between the semantic descriptions. Presumably, the extent of identification as measured in this manner was *not* under voluntary control by the *Ss*. The two adjective descriptions were similarly contrasted and a "distance" score computed for each comparison. Because of the apposition and manifestness of the adjective characterizations, it may be presumed that identification as measured in this way is under appreciable control by the participating *Ss*.

The anticipation with which this study began was that some individuals would show a closeness of ideal self and like-sexed parent descriptions as measured by the direct adjec-

tive check-list approach but marked dissimilarity of ideal self and parent as measured by the semantic differential. It was hypothesized that *Ss* with this pattern of identification would prove to be unusually suppressive or repressive individuals, the latter characteristics being measured independently.

The intended analysis never proved worthwhile, for when the two distributions of discrepancy scores were correlated, the r obtained, based upon an N of 92, proved to be .73. This correlation, when corrected for attenuation became .94. A corrected correlation of this magnitude means that the disguised and manifest means of evaluating identification were functionally equivalent in the way they ordered individuals. Information accessible by the one approach was completely accessible also by the other means.

For the writer, this finding cannot be ascribed to peculiarities of the present sample because it seems likely that in a college sample of this size, a sufficient number of repressive and misidentified *Ss* exists to introduce discrepancy between the check list and differential indices.

Some details of the data analysis should be spelled out. The group of 112 *Ss* was partitioned into four subgroups on the basis of different kinds of response sets observed in the data (e.g., dichotomous rather than trichotomous response to the adjective check list, use of a six-point instead of a seven-point continuum in rating the semantic differential scales, etc.). Analysis of variance indicated three of these subgroups fairly could be combined and the result cited above is based upon this larger sample because of the greater stability afforded the primary result.

Each of the subgroups was analyzed separately, with respect to the correlation of the two measures of identification. The correlations between the semantic differential and adjective check-list measures, corrected for attenuation, ranged from .67 to 1.20 (this latter excess is presumably due to a fortuitously low estimate of the reliability of the semantic differential measure), so that the present findings appear to be generalizable over the various response sets observed. In the four independent comparisons of the subgroups, the split-half reliability of the semantic differential identification measure was always less than that of the adjective check-list measure. For the N of 92, the semantic differential and adjective check-list split-half reliabilities are .70 and .86, respectively.

Although the semantic differential justifiably will continue to enjoy widespread usage, especially with reference to problems of meaning and phenomenology, it would appear from the results of the present study that the technique has certain limitations when it is used to develop *derived* measures of psychological functions. In this kind of research context, the semantic differential may be a rather complicated way of developing a measure more readily and reliably secured by other means.

Received July 3, 1957.

References

1. Cronbach, L. J., & Gleser, Goldine C. Assessing similarity between profiles. *Psychol. Bull.*, 1953, **50**, 456–473.
2. Osgood, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, **49**, 192–237.
3. Osgood, C. E., & Suci, G. J. A measure of relation determined by both mean difference and profile information. *Psychol. Bull.*, 1952, **49**, 251–262.
4. Sanford, N. The dynamics of identification. *Psychol. Rev.*, 1955, **62**, 106–118.
5. Stoke, S. M. An inquiry into the concept of identification. *J. gen. Psychol.*, 1950, **76**, 163–189.

Journal of Consulting Psychology
Vol. 22, No. 3, 1958

Varieties of Response Sets[1]

William E. Broen, Jr.
University of Kansas

and Robert D. Wirt
University of Minnesota

The importance of response sets has long been recognized in animal experimentation. There has recently been an increasing awareness of the usefulness of measures of response sets in psychodiagnostic work. Whether response set scales are seen as possible potential measures of valid variance as Cronbach would think (1, 2), or as measures of something which must be corrected for or suppressed as Fricke suggests (4), all authors seem to agree that understanding of response sets and the development of scales to measure them are necessary for maximal validity in psychological testing. The primary need in this area seems to be more knowledge of the interrelations of response sets.

Although it has been assumed that there are many types of response sets and many measures of supposedly different response sets have been constructed, there is still no accurate empirical knowledge as to whether or not all of these scales measure the same broad general response set, or whether there are a few specific response sets or many unrelated response sets.

Method

It was believed that the above questions could best be answered by factor analysis. Because the completeness of the factor-analytic results would depend upon the completeness of the coverage of the response-set area, the literature was first searched for previously hypothesized varieties of response sets. Measures for 11 possible response sets were then developed; these were: tendency to agree, tendency to disagree, evasiveness, tendency to go to the extreme in agreeing or disagreeing, carefulness, speed, tendency to guess, tendency to check many items, tendency to list many items, tendency to spread judgments on a rating scale, and tendency to rate to the left or right of center. A previously studied set to favor certain answer positions in multiple choice situations was not included because of evidence which cast doubts on existence of such a set (2). Response-set consistency (i.e., stability of set demonstrated by testing with another scale) was also measured in the present study for the tendencies to agree, disagree, and evade, and for the tendency to spread judgments on a rating scale.

An attempt was made to keep the possible influence of item content on the test responses at a minimum.[2] This was done in part by minimizing content, e.g., part of the rating scale data was gathered under directions stating that it was a study in ESP and the subjects were to try to guess the pattern of a person who had made ratings three months previously. In those sections where content was necessary, an effort was made to minimize its influence by using "neutral" content as much as possible. For instance, the content for

[1] We wish to express our gratitude to Sarah Holbrook, senior psychologist, Minneapolis Public Schools; Walter Chapman, principal of Marshall High School, Minneapolis; Frances Anderson, teacher of senior problems at Marshall High School and her students whose generous cooperation made possible the collection of data reported in this study. We also wish to acknowledge the assistance of Chú Cheng-Ping in the analysis of the data.

[2] A copy of the Questionnaire has been deposited with the American Documentation Institute. Order Document No. 5539, remitting $1.75 for 35 mm. microfilm or $2.50 for 6 by 8 in. photocopies.

measures of tendency to agree, disagree, and evade was Fricke's Set *T*, which consists of those items of his Opinion Attitude and Interest Survey (5), which were marked as true by 40 to 60 per cent of each of his criterion groups.

The scales were then administered to three complete classes of high school seniors, all about 17 years old. Fifteen response-set scores (11 response sets, four reliability scores) were obtained for each of the 69 subjects.

Each of the 15 response-set variables was then correlated with each of the others (product-moment correlations) and the resulting matrix was factor-analyzed by the centroid method. In order to achieve more nearly accurate final loadings, the extraction of factors was continued to one factor beyond those indicated to be significant by the application of Humphries' Rule and Tucker's Phi, and all factors were carried in rotation. The nonsignificant factor was "residualized."

Results

Three of the four response sets checked for reliability showed consistency in retesting with different scales. The two measures of evasiveness correlated + .66; those for tendency to agree correlated + .56; and those for tendency to disagree correlated + .58. These correlations are significant beyond the .01 level. There was no intertest consistency for the tendency to spread ratings on a rating scale. The correlation between two measures of this previously hypothesized response set was + .03. A fifth pair of measures which could be considered as a measure of response-set consistency were the variables: tendency to check many adjectives as descriptive of the self, and tendency to list many adjectives in describing the self. These two variables were also significantly correlated ($r = + .41$).

Additional results included the finding that the tendency to go to the extreme in agreeing or disagreeing with personality statements about the self was correlated significantly with both measures of acquiescence (+ .41 and + .45) but was not correlated with the measures of tendency to disagree and that speed was not significantly correlated with carelessness. In fact the correlation, although insignificant (+ .13), tended to show that those who were most careless tended to work more slowly.

Table 1

Rotated Factor Loadings

Response set	Factor I	Factor II	Factor III	h^2
1. Tendency to go to extremes on "neutral" questions	.61	.11	−.07	.39
2. Tendency to agree with "neutral" items	.85	−.25	.03	.79
3. Tendency to disagree with "neutral" items	−.46	.71	−.07	.72
4. Tendency to be undecided on "neutral" items	−.65	−.51	.02	.68
5. Tendency to agree with "neutral" items	.71	.03	−.01	.51
6. Tendency to disagree with "neutral" items	−.31	.82	.00	.77
7. Tendency to be undecided on "neutral" items	−.33	−.78	.02	.72
8. Tendency to guess (small score = guessing)	.13	−.12	.03	.03
9. Tendency to carelessness (high score = carelessness)	.07	−.19	−.09	.05
10. (Time) Tendency to work quickly (high score = more time)	−.07	.34	−.15	.14
11. Tendency to check many adjectives descriptive of self	−.02	−.10	.53	.29
12. Tendency to spread ratings	.05	.01	−.07	.01
13. Tendency to list many adjectives descriptive of self	−.18	−.03	.65	.46
14. Tendency to spread ratings	.13	.03	.36	.15
15. Tendency to rate to the left of center on rating scales or, tending to rate "good"	−.05	.03	.16	.03

Application of Humphries' Rule and Tucker's Phi indicated that three factors were necessary to account for the significant portion of the common variance. The intercorrelations of the factors are as follows: Factors I and II, − .16; I and III, + .01; II and III, .00. The loadings for the 15 response-set variables on these factors are listed in Table 1.

Discussion

The results of this study corroborate those of other studies which show that test results are determined in part by tendencies to respond to items in certain ways regardless of

the specific content of those items. Response-set stability during two testings with different scales was demonstrated for the tendencies to agree, disagree, evade, and the tendency to list or check many items in an adjective check list. The data contraindicated the existence of a consistent response set of spreading vs. narrowing of ratings on rating scales or a speed vs. carefulness set.

A study of the factor loadings for Factor I indicates that it is a dimension with tendency to agree that a statement is true, to make a decision and state it on the one hand; versus the tendencies to deny, to evade, or to be undecided on the other.

The second factor seems to be a dimension of disagreeing with assertions vs. evading the question or being undecided. These data indicate that tendency to disagree is not simply the opposite of an acquiescence set and both cannot be measured on a single scale. Tendency to disagree can be either a tendency to evade the committing of oneself to a particular assertion, and thus much like indecision, and opposed to the making of a positive statement; or it can be the making of a definite negative assertion, and thus be the opposite of evasion or indecision. The first factor can then be seen as a committing onself vs. evasiveness factor (note that tendency to go to the extremes—either very true or very false—has a loading of + .61), and the second factor can be seen as a positive denial vs. indecision factor. It is to be noted that although two scales are needed to measure agreement and denial, these two scales in combination define all points on the plane in which tendency to go to the extreme and indecision tendency are also located.

Factor III seems to indicate a rather specific tendency to list or check many adjectives.

This three-factor analysis does not, of course, indicate that all response sets can be measured by three scales. Rather, one of its main conclusions is the specificity of response sets. Two scales can be used to measure what was previously thought of as the separate denial, acquiescence, tendency to go to the extremes, and evasiveness sets, and a third can measure reasonably well both the hypothesized separate sets of number of adjectives listed and number checked. However, the low commonality on the other response sets shows a need for further scales to measure them. Other studies have reported consistent sets to work for speed, to guess, and to place check marks in certain portions of rating scales (see Cronbach's review of the literature [2]). But each of these demonstrated response sets received very low loadings on each of the factors isolated in our study and each had low intercorrelation with all the other response-set measures. Additional scales would be needed to measure them.

There is a need for measures of, and further research on, all the varieties of response sets. The feeling that all that is necessary is the study of only those response sets which can enter into the particular type of scale used in any particular diagnostic setting stems from the common, but erroneous, view that response sets are only something to be avoided or corrected away. For instance, Fricke (4) stresses only the usefulness of measures of response sets as suppressor scales and goes on to say "Optimal weighting of suppressor and uncorrected scores should probably result in *no correlation between the suppressor and corrected scores*" (the italics are Fricke's [4], p. 163). Statements such as the above are accurate only in those cases where there is no correlation between the response set itself and the criterion for which prediction is being attempted. If there is a correlation between a response set and a criterion, then either the "suppression" of that set *or the neglect to use tests in which that response set can operate* may lead to lower validities.

The clinical meaning of such response sets is not altogether clear. However, other work (5) has indicated, as we would emphasize, that these sets are extremely important in psychodiagnostics. We might speculate, for example, that our Factor III is related to energy level or to extroversion, and personality scales measuring these dimensions should then account for such a factor. Or, it is possible that there are development trends in response sets which, as we suggested in an earlier study (6), distort the results of personality questionnaires. Our Factor III, for example, may primarily reflect a phenomenon of adolescence. The same may be said for other factors. In an unpublished study by the authors,

using college sophomores, we found, for example, high loadings on a common factor for: tendency to say "true" to MMPI items, the Taylor Manifest Anxiety Scale score, the MMPI scale *F* (all high negative loadings) with MMPI scales *L*, *Hy*-subtle, and *K* (all high positive loadings). Looked at another way, 50% of the variance of the *A* scale can be accounted for by certain "test-taking attitude" scales. We believe these relate to response sets and are operative whether the material is relatively "neutral" or "meaningful" clinically. It would seem to us naive to think that clinically important personality syndromes would not include certain response sets as relevant *clinical* features. For these reasons, further research in this area is needed to find out what are the relations of response sets to the various criteria in which we are interested.

Summary

The consistency and interrelations of 11 hypothesized varieties of response sets were studied. A factor analysis of the intercorrelation of these response sets yielded three factors: a committing oneself vs. an evasiveness factor, a positive denial vs. an indecision factor, and a tendency to list or check many adjectives factor. The existence of clinically relevant response-set factors was hypothesized. The uses of measures of the varieties of response sets and the need for further research with them were discussed.

Received July 10, 1957.

References

1. Cronbach, L. J. Response sets and test validity. *Educ. psychol. Measmt*, 1946, **6**, 475–494.
2. Cronbach, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt*, 1950, **10**, 3–31.
3. Fricke, B. G. *The Opinion Attitude and Interest Survey*. Minneapolis: Investor Diversified Services, 1955.
4. Fricke, B. G. Response set as a suppressor variable in the OAIS and the MMPI. *J. consult. Psychol.*, 1956, **20**, 161–169.
5. Meehl, P. E., & Hathaway, S. R. The K factor as a suppressor variable in the MMPI. *J. appl. Psychol.*, 1946, **30**, 525–564.
6. Wirt, R. D., & Broen, W. E., Jr. The relation of the Children's Manifest Anxiety Scale to the concept of anxiety as used in the clinic. *J. consult. Psychol.*, 1956, **20**, 482.

PSYCHOLOGICAL TEST REVIEWS

New Tests

Berdie, Ralph F., & Layton, Wilbur R. *Minnesota Counseling Inventory*. Grades 9–12. 1 form. Untimed, (50) min. IBM, Testscor, or hand scoring. Inventory booklet ($3.50 per 25); IBM answer sheet and profile sheet ($3.60 per 50); Hankes answer sheet ($1.90 per 50); profile sheet only ($1.90 per 50); manual, pp. 27, with hand-scoring keys (60¢), with machine-scoring keys (70¢); specimen set (75¢). New York: Psychological Corp., 1953, 1957.

The Minnesota Counseling Inventory is an instrument for counseling high school students, based on two earlier questionnaires. From the Minnesota Personality Scale (1941), three scales were taken: Family Relationships (*FR*), Social Relationships (*SR*), and Emotional Stability (*ES*). Five MMPI scales represented by selected and sometimes edited items are *L, Pd, Sc, D,* and *Si,* which in the new Inventory are designated, respectively, as Validity (*V*), Conformity (*C*), Adjustment to Reality (*R*), Mood (*M*), and Leadership (*L*). Forty-four of the 355 items are not presently scored but represent potentialities for the development of new scales.

The extensive data given in the Manual reveal both favorable and unfavorable features of the development of the Inventory. Both split-half and retest reliabilities are mainly satisfactory, ranging from .56 to .95 in two-grade, one-sex groups, with the typical coefficient in the middle .80's. Standard score norms, given separately for boys and girls and for grades 9–10 and 11–12, are based on adequate methods and numbers. The manual's suggestions for interpretation are explicit and appropriately modest.

One weakness is revealed by the data on validities and intercorrelations. Validities of entire scales are reported in terms of the CRs of differences between the scores of random samples of pupils and of extreme cases nominated by teachers and school administrators. There is no indication of item analysis against a criterion. Some scales, such as *SR,* differentiate well, but others, such as *R* (*Sc*), fail to do so. The tables of intercorrelations show that most *r*'s run from .35 to .60, but that the correlations between *SR* and *L,* and between *ES* and *R,* are as high as the scales' reliabilities. The authors might well have refined the scales further by item analysis, or at least might have warned users that the evidence points to five interpretable scores, not seven. Even so, the Inventory is a serviceable instrument for counselors whose knowledge of test construction permits them to be aware of its limitations.—*L. F. S.*

Lamke, Tom A., & Nelson, M. J. *Henmon-Nelson Tests of Mental Ability*. (Rev. ed.) Grades 3–12. 3 levels, 2 forms at each level. 30 (40) min. Test booklet, each level ($3.00 per 35), with manual, pp. 24, and record sheet; specimen set (42¢). Boston: Houghton Mifflin, 1957.

The widely used Henmon-Nelson, which first appeared in 1931, has been given a thorough revision and restandardization. The test is now published in three levels, for grades 3–6, 6–9, and 9–12. Each level contains 90 multiple-choice items in spiral-omnibus form, with several types of verbal items predominating; there are some numerical items and figure analogies. A fourth level, for college students, is in preparation. The test booklets are consumable and are scored rapidly by the use of a carbon-paper backing that eliminates the need for a key. Norms for each level are provided in terms of percentile ranks in grade, grade equivalents, MAs and deviation IQs.

The content of the revision is based on an extensive item analysis of the older tests, and the equivalence of the Forms A and B is well established. The separate manuals for the three levels give impressively favorable data about reliability, congruent validity with other tests, and concurrent and predictive validity with achievement measures. Reliabilities generally exceed .90, correlations with other intelligence tests approximate .75, and correlations with achievement are high, typically .70 with relevant tests and .60 with school grades. The standardization is a model of precise method, using a predetermined sample of 250 classrooms stratified by state and size of school system. The manual closes with a persuasive argument for the use of omnibus tests, rather than of tests with factor scores, for the prediction of school achievement.

In all, the revised Henmon-Nelson impresses the reviewer as a scholarly example of the best in test construction, and as a remarkably efficient instrument for its length.—*L. F. S.*

Personal Adjustment: An Approach through the Study of Healthy Personality

By SIDNEY M. JOURARD, *Ph.D., Associate Professor of Psychology, University of Florida.* Written from a positive point of view, this unique work discusses the factors determining the healthy, adjusted personality. The author outlines each aspect of personality and compares the normal, unhealthy and healthy versions of the trait in question. The sections on expositions of interpersonal relations, love, the self-structure, and conscience are new to the field.

A pre-publication reader comments: "I think it is indeed a new approach to mental hygiene and a far better one than what we now have available." January, 1958, 462 pages, $5.50

INTRODUCTORY CLINICAL PSYCHOLOGY

By SOL L. GARFIELD, *University of Nebraska Medical School.* Stressing diagnosis, psychotherapy and research, this book presents a study of the development and current role of clinical psychology, both as a science and as a profession. *1957, 469 pages, $6.00*

DIFFERENTIAL PSYCHOLOGY: Individual and Group Differences in Behavior, Third Edition

By ANNE ANASTASI, *Fordham University.* Thoroughly revised to include recent American and European research in biology, anthropology and sociology, the third edition of this authoritative volume contains much significant new material. *1958, 664 pages, $7.50*

The Macmillan Company

60 FIFTH AVENUE, NEW YORK 11, N.Y.

A NEW PUBLICATION

VOL. I, PART I — JAN–MARCH 1958

LANGUAGE AND SPEECH

***Edited by* D. B. FRY**

This new quarterly journal is devoted principally to experimental research in the field of language and speech. The first issue is now available and contains the following papers:

KATHERINE S. HARRIS (*Haskins Laboratories, New York*). Cues for the discrimination of American English fricatives in spoken syllables.

G. HERDAN (*University of Bristol*). The relation between the functional burdening of phonemes and their frequency of occurrence.

A. R. LURIA (*University of Moscow*). Brain disorders and language analysis.

D. B. FRY AND P. DENES (*University College, London*). The solution of some fundamental problems in mechanical speech recognition.

FRIEDA GOLDMAN-EISLER (*University College, London*). Speech analysis and mental processes.

Annual Subscription: £4 (or $11.50).

Subscriptions should be sent to the publishers:

Robert Draper Ltd.
Kerbihan House, 85 Udney Park Road, Teddington, Middlesex, England.

Contributions should be addressed to:

D. B. Fry.
University College, Gower Street, London W.C.1.

Important Evergreen Psychology Books

$1 and up

The same fine soft-cover editions used in college psychology courses throughout the country. Now at your bookstore.

Masochism in Modern Man by *Theodor Reik*. $1.95

The Biosocial Nature of Man by *Ashley Montagu*. $1.25

A Study of Interpersonal Relations. New Contributions to Psychiatry, edited by *Patrick Mullahy*. $2.45

Listening with the Third Ear. The Inner Experience of a Psychoanalyst by *Theodor Reik*. $1.95

Psychoanalysis: Evolution and Development. A Review of Theory and Therapy, by *Clara Thompson* and *Patrick Mullahy*. $1.75

Democracy and Dictatorship. An incisive study of the psychological and social conditions that produce each, by *Zevedei Barbu*. $1.45

Black Anger. A unique psychoanalytic biography of a native African medicine man, by *Wulf Sachs*. $1.75

Jung's Psychology and Its Social Meaning by *Ira Progoff*. $1.75

Education and Human Relations by *Ashley Montagu*. $1.45

Oedipus: Myth and Complex. A review of all major psychoanalytic theories by *Patrick Mullahy*. $1.95

The Forgotten Language. An Introduction to the Understanding of Dreams, Fairy-tales and Myths, by *Erich Fromm*. $1.75

Specialized Techniques in Psychotherapy. Edited by *Gustav Bychowski, M.D.* and *Louise Despert, M.D.* $1.95

Rebel Without a Cause. The hypnoanalysis of a criminal psychopath, by *Robert M. Lindner*. $1.75

The Marquis de Sade. An essay by *Simone de Beauvoir*, with selections from De Sade's writings. $1.75

SEND FOR FREE DESCRIPTIVE CATALOG

grove press
795 Broadway, New York 3

PSYCHIATRY

Journal for the Study of Interpersonal Processes

CONTENTS OF NO. 2, VOL. 21, MAY, 1958

Vilhelm Aubert: *Legal Justice and Mental Health*

Stanley H. Eldred and Douglas B. Price: *A Linguistic Evaluation of Feeling States in Psychotherapy*

Montague Ullman: *Dreams and the Therapeutic Process*

Seymour Fisher and David Mendell: *The Spread of Psychotherapeutic Effects from the Patient to his Family Group*

Helm Stierlin: *Contrasting Attitudes Toward the Psychoses in Europe and in the United States*

Viola W. Bernard: *School Desegregation—Some Psychiatric Implications*

Earle Silber, Stewart E. Perry, and Donald A. Bloch: *Patterns of Parent-Child Interaction in a Disaster*

Ednita P. Bernabeu: *The Effects of Severe Crippling on the Development of a Group of Children*

Rae Shifrin Sternberg, Jean Chapman, and David Shakow: *Psychotherapy Research and the Problem of Intrusions on Privacy*

Lyman C. Wynne, Irving M. Ryckoff, Juliana Day, and Stanley I. Hirsch: *Pseudo-mutuality in the Family Relations of Schizophrenics*

Editorial Notes

Brief Communications

Book Reviews

Published in February, May, August, and November by the

William Alanson White Psychiatric Foundation

1703 Rhode Island Ave., N. W., Washington 6, D. C.

Subscription price $10.00 for 1958; foreign $10.80

Single copy $2.50; foreign $2.70

Institute: July 14–25 at:

Wiesbaden Germany

for the

KAHN TEST

OF SYMBOL ARRANGEMENT

Participants:

DR. FREDERICK KRUSE

Administration of the Kahn Test: K. Hartmann, Dipl. Psych.; *Scoring:* J. Gamble, M.A.; *Evaluation of Children with the Kahn Test:* A. Eibel, Dipl. Psych., *Chief*, Wiesbaden Institut Fuer Entwicklungshilfe; also E. Wienhoff, Dipl. Psych.; *Psychodiagnosis:* R. Frostick, M.D.; A. Turner, M.D.; and others. *Also a special session on the use of the Kahn Test in the prognosis of therapy.*

Reservations: Dr. K. Hartmann, Rhein St. 11, Wiesbaden, Germany.

FREE LITERATURE AND REPRINTS ON THE KAHN TEST OF SYMBOL ARRANGEMENT are available from Professor R. B. Ammons, Box 1441, Missoula, Montana, U.S.A.

DEVEREUX SERVES

A full educational program, extending from kindergarten through the junior college level, is offered at Devereux for emotionally disturbed, as well as slow-learning children.

Instruction is highly individualized and planned to take advantage of the specialized learning techniques which are most effective for these children.

The curricula of the Devereux units include a solid academic program for college preparatory work, a vocational and commercial program and programs of artisanship or craftsmanship, stressing the use of special tools and skills, either in preparation for admission to a vocational or trade school or as a means of attaining self-confidence and personality development.

Professional inquiries should be addressed to John M. Barclay, Director of Development, Devereux Schools, Devon, Pennsylvania; summer residents address Keith A. Seaton, Registrar, Devereux Schools in California, Santa Barbara, California.

THE DEVEREUX FOUNDATION

A non-profit corporation — *Founded 1912*

Santa Barbara, California — *Devon, Pennsylvania*

SCHOOLS
COMMUNITIES
CAMPS
TRAINING
RESEARCH

Professional
Associate Directors
Charles M. Campbell, Jr., M.D.
Michael S. Dorn, Ph.D.
Paul A. Henry, S.T.D.
J. Clifford Scott, M.D.